

# Magnification bias in galaxy surveys with complex sample selection functions

Maximilian von Wietersheim-Kramsta,<sup>1</sup>★ Benjamin Joachimi,<sup>1</sup> Jan Luca van den Busch,<sup>2</sup>  
Catherine Heymans,<sup>2,3</sup> Hendrik Hildebrandt,<sup>2</sup> Marika Asgari,<sup>3</sup> Tilman Tr’oster,<sup>3</sup> Sandra Unruh<sup>4</sup>  
and Angus H. Wright<sup>2</sup>

<sup>1</sup>Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, UK

<sup>2</sup>Faculty of Physics and Astronomy, Astronomical Institute (AIRUB), Ruhr University Bochum, German Centre for Cosmological Lensing, D-44780 Bochum, Germany

<sup>3</sup>Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, UK

<sup>4</sup>Argelander-Institut für Astronomie, German Centre for Cosmological Lensing, Auf dem Hügel 71, D-53121 Bonn, Germany

Accepted 2021 March 31. Received 2021 March 23; in original form 2021 January 13

## ABSTRACT

Gravitational lensing magnification modifies the observed spatial distribution of galaxies and can severely bias cosmological probes of large-scale structure if not accurately modelled. Standard approaches to modelling this magnification bias may not be applicable in practice as many galaxy samples have complex, often implicit, selection functions. We propose and test a procedure to quantify the magnification bias induced in clustering and galaxy–galaxy lensing (GGL) signals in galaxy samples subject to a selection function beyond a simple flux limit. The method employs realistic mock data to calibrate an effective luminosity function slope,  $\alpha_{\text{obs}}$ , from observed galaxy counts that can then be used with the standard formalism. We demonstrate this method for two galaxy samples derived from the Baryon Oscillation Spectroscopic Survey (BOSS) in the redshift ranges  $0.2 < z \leq 0.5$  and  $0.5 < z \leq 0.75$ , complemented by mock data built from the MICE2 simulation. We obtain  $\alpha_{\text{obs}} = 1.93 \pm 0.05$  and  $\alpha_{\text{obs}} = 2.62 \pm 0.28$  for the two BOSS samples. For BOSS-like lenses, we forecast a contribution of the magnification bias to the GGL signal between the multipole moments,  $\ell$ , of 100 and 4600 with a cumulative signal-to-noise ratio between 0.1 and 1.1 for sources from the Kilo-Degree Survey (KiDS), between 0.4 and 2.0 for sources from the Hyper Suprime-Cam survey (HSC), and between 0.3 and 2.8 for ESA Euclid-like source samples. These contributions are significant enough to require explicit modelling in future analyses of these and similar surveys. Our code is publicly available within the MAGBET module (<https://github.com/mwiet/MAGBET>).

**Key words:** gravitational lensing: weak – methods: data analysis – methods: observational.

## 1 INTRODUCTION

Over the last few decades, weak gravitational lensing has become a powerful tool to directly measure the matter distribution of the late Universe, while allowing for the inference of the cosmological parameters that govern it. Surveys, such as the currently ongoing Kilo Degree Survey<sup>1</sup> (KiDS; Kuijken et al. 2015), the Dark Energy Survey<sup>2</sup> (DES; Flaugher et al. 2015), the Hyper Suprime-Cam Subaru Strategic Program<sup>3</sup> (HSC SSP; Aihara et al. 2018), have become increasingly limited by systematics rather than statistics as ever-growing sample sizes reduce uncertainties. The impact of the systematics will become even more exaggerated for the next generation of surveys, e.g. *Euclid*<sup>4</sup> (Laureijs et al. 2011), the Vera C.

Rubin Observatory Legacy Survey of Space and Time<sup>5</sup> (LSST; Abell et al. 2009), and the Nancy Grace Roman Space Telescope<sup>6</sup> (also known as WFIRST; Spergel et al. 2015). For this reason, recent efforts have focused on improving our physical understanding of often neglected phenomena that can influence cosmological parameter inference based on shear and clustering measurements. These effects include intrinsic galaxy alignments (Kiessling et al. 2015; Kirk et al. 2015; Troxel & Ishak 2015) and magnification (Hildebrandt, Van Waerbeke & Erben 2009; Duncan et al. 2014; Hildebrandt 2016; Thiele, Duncan & Alonso 2020; Unruh et al. 2020). In this paper, we will focus on the magnification effects.

While the magnification due to gravitational lensing partially manifests itself as a change in the angular diameter of an object, it also changes the observed solid angle of a field with respect to the intrinsic solid angle. This can affect the observed galaxy counts and their fluxes, leading to magnification effects that have been

\* E-mail: [maximilian.wietersheim-kramsta.19@ucl.ac.uk](mailto:maximilian.wietersheim-kramsta.19@ucl.ac.uk)

<sup>1</sup><https://kids.strw.leidenuniv.nl>

<sup>2</sup><https://www.darkenergysurvey.org>

<sup>3</sup><https://hsc.mtk.nao.ac.jp/ssp>

<sup>4</sup><https://www.euclid-ec.org>

<sup>5</sup><https://www.lsst.org>

<sup>6</sup><https://roman.gsfc.nasa.gov>

detected in the past by Chiu et al. (2016) and Garcia-Fernandez et al. (2018). It is important to note that this affects the counts of source galaxies and lens galaxies, such that the magnification due to large-scale structure can also change the shear-clustering cross-correlations (galaxy–galaxy lensing; GGL) and the clustering measurements (Hui, Gaztañaga & LoVerde 2007; Ziour & Hui 2008; Duncan et al. 2014; Thiele et al. 2020; Unruh et al. 2020). Therefore, if this effect is not accurately modelled in such analyses, a *magnification bias* can be induced. However, we also note that, in the literature and in this paper, the term *magnification bias* is regularly used to refer magnification effects even when they are modelled.

We break down the magnification effect into two separate phenomena: *flux magnification* and *lensing dilution*. The first is caused by an increase/decrease in the flux observed from a source due to gravitational lensing that can push otherwise unobserved galaxies over the flux limit or push galaxies with magnitudes below the flux limit out of the observational window. At the same time, lensing dilution increases/decreases the number of observed sources within a certain area of the sky by (de-)magnifying the solid angle behind the gravitational lens. The magnification effect can be measured directly from changes in the apparent size and magnitude of lensed galaxies (Schmidt et al. 2011) or by comparing the observed galaxy effective radii to the intrinsic radii derived from their surface brightness and stellar velocity dispersion (Huff & Graves 2013). None the less, it is most commonly measured through the bias in the observed number density of sources (Scranton et al. 2005). Since this bias directly contributes to the clustering and GGL signal, we will rely on this approach in our analysis.

The constraining power of weak lensing samples is constantly growing (Troxel & Ishak 2015; Hikage et al. 2019; Asgari et al. 2020) by including additional measurements (Abbott et al. 2018, 2019a, b) and through joint analyses between different surveys like, for example, in the recent joint analysis of KiDS-1000 with BOSS (methodology described in Joachimi et al. 2020 and the results are shown in Heymans et al. 2020 and in Tröster et al. 2020). In all these analyses, the understanding of the systematics is becoming a priority. One potential systematic could appear from unaccounted magnification biases in the clustering signal of a non-flux-limited spectroscopic surveys such as BOSS (Dawson et al. 2012) or DESI<sup>7</sup> (Aghamousa et al. 2016) or colour-selected photometric samples such as DES REDMAGIC (Rozo et al. 2016) or luminous red galaxy samples (Vakili et al. 2020). Thus also biasing the GGL correlations with shear signal from weak lensing surveys.

This paper aims to provide a method for estimation of the magnification bias for surveys that have complex sample selection functions that are not purely flux/magnitude-limited. We use the standard framework for estimating the magnification bias from observables in flux-limited surveys as a basis for the parametrization of a semi-empirical model for non-flux-limited surveys. This model is then tested by comparing the estimates for the magnification bias in BOSS observations (Dawson et al. 2012) to the estimates from MICE2 cosmological simulations. We then use our results to forecast some of the potential biases that could be induced in a joint analysis of KiDS-1000 or HSC Wide with BOSS and a *Euclid*-like survey with a DESI-like survey.

This article is structured in the following manner. In Section 2, the theoretical background is described. In Section 3, we provide an outline and presentation of our methods and simulations. The magnification bias estimates from a BOSS-like galaxy population are presented in Section 4. The forecasts for current and future joint

analyses are found in Section 5. Lastly, we conclude the paper and provide an outlook in Section 6. Appendix A repeats the analysis shown in Section 4 for a magnitude limited galaxy sample.

## 2 THEORETICAL BACKGROUND

### 2.1 Magnification bias for flux-limited surveys

As described in the review by Bartelmann & Schneider (2001), a lensed population of galaxies with a cumulative galaxy count  $N$  at redshift  $z$ , given a flux limit of  $S$ , can be described in terms of the unlensed population,  $N_0$ , as

$$N(> S, z) = \frac{1}{\mu(z)} N_0 \left( > \frac{S}{\mu(z)}, z \right), \quad (1)$$

where  $\mu(z)$  is the magnification for a redshift  $z$ . Here, the  $1/\mu(z)$  factor accounts for the dilution of galaxies due to magnification. The unlensed population has been observationally shown to be similar to a power law in flux (in particular, for faint galaxies) given by

$$N_0(> S, z) = AS^{-\alpha} p_0(z; S), \quad (2)$$

where  $A$  and  $\alpha$  parametrize the power law and  $p_0(z; S)$  is the redshift probability distribution of the galaxies. Taking the ratio of these two populations, assuming that we can approximate the  $\mu(z)$  with the magnification  $\mu$  of a fiducial source at infinity (which should hold mainly at low redshifts; Bartelmann & Schneider 2001) and integrating over redshift, we get the following expression:

$$\frac{N(> S)}{N_0(> S)} = \mu^{\alpha-1}. \quad (3)$$

If  $\alpha \approx 1$ , we can see from equation (3) that the magnification bias would vanish (with slight deviations from this depending on the redshift range). The magnification can be related directly to the local surface density  $\kappa$  in the weak lensing limit ( $|\kappa| \ll 1$ ,  $|\gamma| \ll 1$ ) with  $\mu \approx 1 + 2\kappa$  (Broadhurst & Lehár 1995). Therefore, one can relate  $\kappa$  to the relative difference between the magnified and the unmagnified galaxy populations and the exponent  $\alpha$  of the flux power spectrum with

$$\frac{N(> S) - N_0(> S)}{N_0(> S)} \approx 2(\alpha_\kappa - 1)\kappa, \quad (4)$$

where  $\alpha_\kappa$  is the same as the  $\alpha$  in equation (3) in the weak lensing limit. When analysing samples with a complex selection function, equation (4) does not necessarily apply anymore. None the less, we use the parameter  $\alpha_\kappa$  as an analogue to estimate the magnitude of the magnification bias in a given galaxy sample.

### 2.2 Estimating the magnification bias in flux-limited surveys

By considering equations (1) and (2), and the definition of magnitude as a function of flux, one can derive that  $\alpha_{\text{obs}}$  can be determined from the differential galaxy count  $n(m)$  over a given band magnitude range from  $m$  to  $m + dm$  as follows (Binggeli, Sandage & Tammann 1988; Bartelmann & Schneider 2001; Hildebrandt et al. 2009):

$$\alpha_{\text{obs}}(m) = 2.5 \frac{d \log_{10}[n(m)]}{dm}. \quad (5)$$

One could get the same estimates of  $\alpha_{\text{obs}}(m)$  (at least, for a flux-limited sample) by replacing  $n(m)$  in equation (5) with the cumulative galaxy count distribution. However, here we choose to derive  $\alpha_{\text{obs}}(m)$  from the differential distribution,  $n(m)$ , instead, because we find that it gives more robust estimates when deviating from the flux-limited case. Also, note that sometimes the differential galaxy count distribution is given over flux,  $S$ , instead of magnitude,  $m$ . Then,  $\alpha_{\text{obs}}(S) + 1$  is given by  $-\text{dlog}_{10}[n(S)]/\text{dlog}_{10}(S)$ .

<sup>7</sup><https://www.desi.lbl.gov>

This  $\alpha_{\text{obs}}$  near the faint end of the galaxy population is considered as an *effective* luminosity function slope if it is consistent with the  $\alpha_\kappa$  value given by equation (4). Therefore, by estimating the luminosity function slope,  $\alpha_\kappa$ , through the observed  $\alpha$ , one can estimate the systematic effects that may be introduced to galaxy number counts through the magnification bias, and therefore the systematics affecting the clustering and GGL signals derived from this observable.

### 2.3 Signal modelling

In accordance with the framework outlined in Section 2 of Joachimi et al. (2020) as the methodology for the inference of cosmological parameters from KiDS-1000, we opt to quantify the influence of the magnification bias on cosmology through its contribution to the GGL angular power spectra. These angular power spectra are line-of-sight projections of the three-dimensional matter power spectrum. We express the observable GGL angular power spectrum correlating galaxy positions,  $n$ , and galaxy shapes,  $\epsilon$ , as a linear functional of derived statistics as

$$C_{\text{ne}}^{(ij)}(\ell) = C_{\text{gG}}^{(ij)}(\ell) + C_{\text{gl}}^{(ij)}(\ell) + C_{\text{mG}}^{(ij)}(\ell), \quad (6)$$

where  $i$  is the index for lens galaxy redshift bins,  $j$  is the index of the source galaxy samples, gG stands for the cross-correlation between the lens galaxy distribution, and the source gravitational shear, gl stands for the intrinsic alignment of source galaxies physically close to foreground lenses and mG stands for the correlation between gravitational shear and the lensing-induced magnification bias in the lens sample.  $C_{\text{ga}}^{(ij)}(\ell)$  for  $a \in \{\text{G}, \text{I}\}$  are defined as Limber-approximated line-of-sight projections of the three-dimensional cross-power spectrum between the galaxy and matter distribution,  $P_{\text{gm}}$ , given by (Kaiser 1992; LoVerde & Afshordi 2008)

$$C_{\text{ga}}^{(ij)}(\ell) = \int_0^{\chi_{\text{hor}}} \frac{n_{\text{L}}^{(i)}(\chi) W_{\text{a}}^{(j)}(\chi)}{f_{\text{K}}^2(\chi)} P_{\text{gm}}\left(\frac{\ell + 1/2}{f_{\text{K}}(\chi)}, \chi\right) d\chi, \quad (7)$$

where  $\chi$  is the comoving distance,  $\chi_{\text{hor}}$  is the comoving distance to the horizon,  $n_{\text{L}}^{(i)}$  is the comoving distance distribution of the lens sample  $i$ , and  $f_{\text{K}}$  is the comoving angular diameter distance.  $W_{\text{G}}^{(j)}$  is the weak lensing kernel and is given by

$$W_{\text{G}}^{(j)}(\chi) = \frac{3H_0^2 \Omega_{\text{m}}}{2c^2} \frac{f_{\text{K}}(\chi)}{a(\chi)} \int_{\chi}^{\chi_{\text{hor}}} d\chi' n_{\text{S}}^{(j)}(\chi') \frac{f_{\text{K}}(\chi' - \chi)}{f_{\text{K}}(\chi')}, \quad (8)$$

where  $H_0$  is the Hubble constant,  $\Omega_{\text{m}}$  is the matter density parameter,  $c$  is the speed of light,  $a$  is the scale factor, and  $n_{\text{S}}^{(j)}$  is the comoving distance distribution of the source sample  $j$ .  $W_{\text{I}}^{(j)}$  is the intrinsic alignment (IA) kernel. Here, we choose an IA kernel in accordance with the non-linear alignment model (Bridle & King 2007) given by

$$W_{\text{I}}^{(j)}(\chi) = -A_{\text{IA}} \frac{C_1 \rho_{\text{cr}} \Omega_{\text{m}}}{D(a[\chi])} n_{\text{S}}^{(j)}(\chi), \quad (9)$$

where  $A_{\text{IA}}$  is the IA amplitude,  $z_{\text{pivot}}$  is an arbitrary pivot that is set to 0.3 in line with previous IA analyses (Joachimi et al. 2011),  $C_1$  denotes a normalization constant,  $\rho_{\text{cr}}$  is the critical density,  $D$  is the linear growth factor normalized to unity at the present day. We normalize the IA kernel by setting  $C_1 \rho_{\text{cr}} \approx 0.0134$ , i.e.  $C_1 = 5 \times 10^{-14} (h^2 \text{ M}_{\odot}/\text{Mpc}^{-3})^{-2}$ , in accordance with the value from Hirata & Seljak 2004 and Bridle & King 2007 that is set using the galaxy ellipticity measurements from SuperCOSMOS (Hambly et al. 2001; Brown et al. 2002).

The magnification term in equation (6) is modelled as

$$C_{\text{mG}}^{(ij)}(\ell) = 2(\alpha_{\text{obs}}^{(i)} - 1) C_{\text{gG}}^{(ij)}(\ell), \quad (10)$$

where  $i$  again indexes lens galaxy samples,  $j$  indexes source samples, mG stands for the lensing-induced magnification bias in the lens sample, and GG stands for shear-shear correlation signal.  $C_{\text{GG}}^{(ij)}(\ell)$  is defined as the cosmic shear angular power spectrum purely from gravitational lensing effects, i.e. without any intrinsic alignment signals, and is given by

$$C_{\text{GG}}^{(ij)}(\ell) = \int_0^{\chi_{\text{hor}}} \frac{W_{\text{G}}^{(i)}(\chi) W_{\text{G}}^{(j)}(\chi)}{f_{\text{K}}^2(\chi)} P_{\text{m, nl}}\left(\frac{\ell + 1/2}{f_{\text{K}}(\chi)}, z(\chi)\right) d\chi, \quad (11)$$

where  $P_{\text{m, nl}}$  is the non-linear matter power spectrum. This power spectrum is computed with a non-perturbative model using HM-CODE (Mead et al. 2015, 2016) integrated within CAMB<sup>8</sup> (Lewis, Challinor & Lasenby 2000; Lewis & Bridle 2002; Howlett et al. 2012). HM-CODE incorporates baryonic feedback in its halo modelling approach. We solely parametrize the baryonic feedback model using one free parameter,  $A_{\text{bary}}$ , in line with Hildebrandt et al. (2017). The non-linear matter power spectrum  $P_{\text{m, nl}}$  is also used to compute the cross-power spectrum between the galaxy and matter distribution  $P_{\text{gm}}$  used in equation (7) as in the analysis shown in Joachimi et al. (2020).

## 3 METHODOLOGY

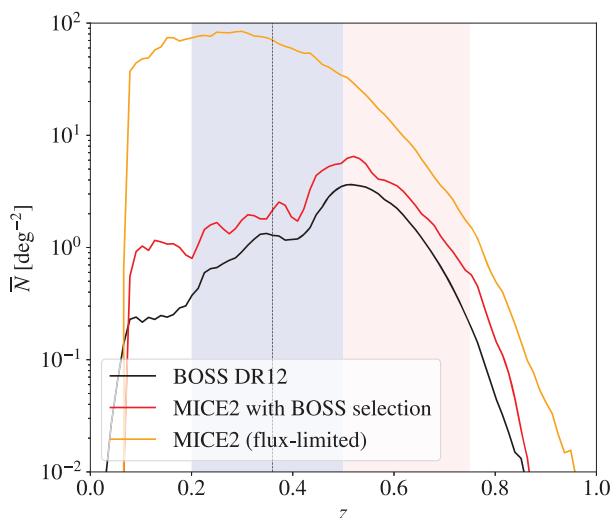
The method outlined in this paper aims to provide an accurate estimate of the *effective* luminosity function slope,  $\alpha$ , of a galaxy sample with a complex sample selection. This estimate can be used to quantify the magnification bias in clustering and GGL-lensing analyses. To achieve this, we rely on realistic weak lensing simulations to calibrate the  $\alpha_{\text{obs}}$  estimate from observables, based on equation (5), such that it agrees with the value of  $\alpha_\kappa$  derived from unobservable quantities using equation (4). The procedure gives a magnitude range that yields the most optimal  $\alpha_{\text{obs}}$  value. This value is used to estimate  $\alpha_{\text{obs}}$  from observations. If the simulations are accurate,  $\alpha_{\text{obs}}$  should agree with the underlying  $\alpha_\kappa$  even though it cannot be directly measured.

### 3.1 BOSS DR12 data

We develop our method using lens samples derived from the Sloan Digital Sky Survey (SDSS)-III BOSS (Eisenstein et al. 2011; Dawson et al. 2012). BOSS is a spectroscopic survey with a complex sample selection function that is commonly used in cosmological analyses of galaxy clustering and GGL (Alam et al. 2017; Beutler et al. 2017; Sánchez et al. 2017; Tröster et al. 2019; Speagle et al. 2019; Heymans et al. 2020). For more details about the nature of the galaxy selection process, see Alam et al. (2015). A lens galaxy sample selected in such a way could be introducing a substantial magnification bias in any analysis, while its complexity does not allow to measure it with current means. For the BOSS sample, the bias becomes even more important to model, because it is commonly used in GGL analysis with the source galaxy samples of weak lensing surveys whose footprint significantly overlaps with the BOSS footprint.

For this work, we use the photometric data from the final data release of BOSS, DR12 (Alam et al. 2015) with the same target selection as in Sánchez et al. (2017). This sample combines the BOSS LOWZ and CMASS galaxy samples to produce a catalogue that covers approximately 9300 deg<sup>2</sup> (Reid et al. 2016). Its normalized redshift distribution can be seen in Fig. 1. The sample is then split into two redshift ranges: ‘zlow’ ( $0.2 < z \leq 0.5$ ) and ‘zhigh’ ( $0.5 < z \leq 0.75$ ). From this photometric data, we use SDSS composite

<sup>8</sup>Code for Anisotropies in the Microwave Background; <https://camb.info>



**Figure 1.** Galaxy counts per unit area on the sky,  $\bar{N}$ , for 100 redshift bins within  $0 < z \leq 1$ . The SDSS DR12 BOSS sample is shown in black, the MICE2 sample with the BOSS selection function in red and the flux-limited MICE2 sample in orange. The blue area marks the domain between  $z = 0.2$  and  $z \leq 0.5$  that defines the zlow bin, while the red area marks the domain of the zhigh bin ( $0.5 < z \leq 0.75$ ). The dashed black horizontal line indicates the boundary between the LOWZ and the CMASS samples within the BOSS DR12 sample at  $z \sim 0.36$ .

model (cmodel) band magnitudes which are defined in Stoughton et al. (2002).

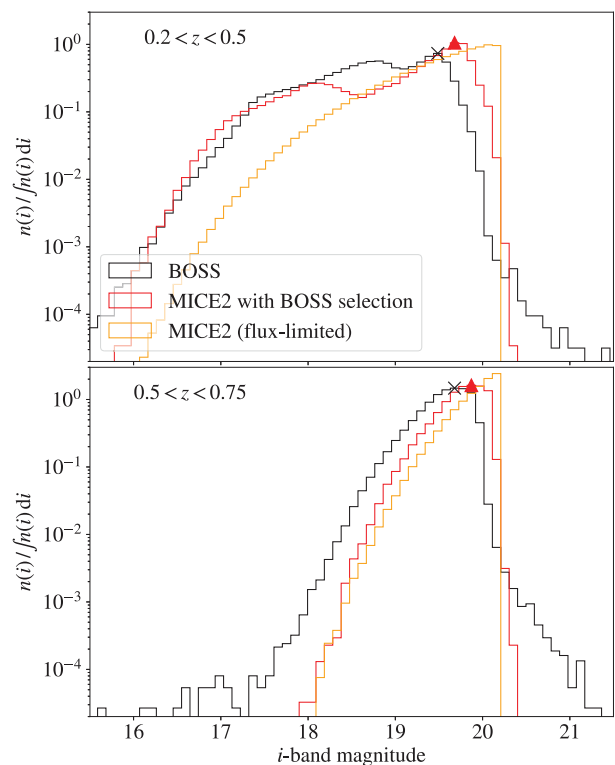
### 3.2 MICE2 simulations

For the analysis discussed in Section 4 and in appendix A, we rely on data sets of simulated galaxies, selected from the MICE2 galaxy mock catalogue (Carretero et al. 2015; Crocce et al. 2015; Fosalba et al. 2015a, b; Hoffmann et al. 2015). This catalogue is based on the MICE dark matter-only simulation, generated from  $7 \times 10^{10}$  particles in a box with a side length of 3 Gpc and assuming a Lambda cold dark matter ( $\Lambda$ CDM) cosmological model with  $\Omega_m = 0.25$ ,  $\Omega_\Lambda = 0.75$ ,  $\Omega_b = 0.044$ , and  $h = 0.7$ . A light cone, spanning  $\sim 5000 \text{ deg}^2$ , is constructed from this simulation box and populated with galaxies up to a redshift of  $z = 1.4$  using a hybrid Halo Occupation Distribution and Halo Abundance Matching technique. Additionally, MICE2 embeds gravitational lensing by providing estimates of the shear components, convergence as well as true and lensed position for each galaxy. MICE2 derives weak lensing properties by constructing all-sky shells in steps of 35 Myr of lookback time (Fosalba et al. 2015b). These are then projected into HEALPIX maps (Gorski et al. 2005) from which the convergence is computed using the Born-approximation (Fosalba et al. 2015b). Therefore, galaxies within the same HEALPIX pixel inherit the same lensing properties which are, due to this limitation, accurate down to scales of 1 arcmin. We compute the magnified galaxy magnitudes according to equation (12) that uses the weak lensing assumption  $|\kappa| \ll 1$  by approximating  $\mu \approx 1 + 2\kappa$ .

We start from this MICE2 input catalogue and apply an evolutionary correction to the provided SDSS  $g'r'i$ -band magnitudes and calculate an additional set of magnitudes

$$m^{\text{mag}} = m^{\text{evo}} - 2.5 \log_{10}(1 + 2\kappa) \quad (12)$$

that factor in magnification, where  $m^{\text{evo}}$  are the evolution corrected MICE2 magnitudes and  $\kappa$  the convergence (see van den Busch et al.



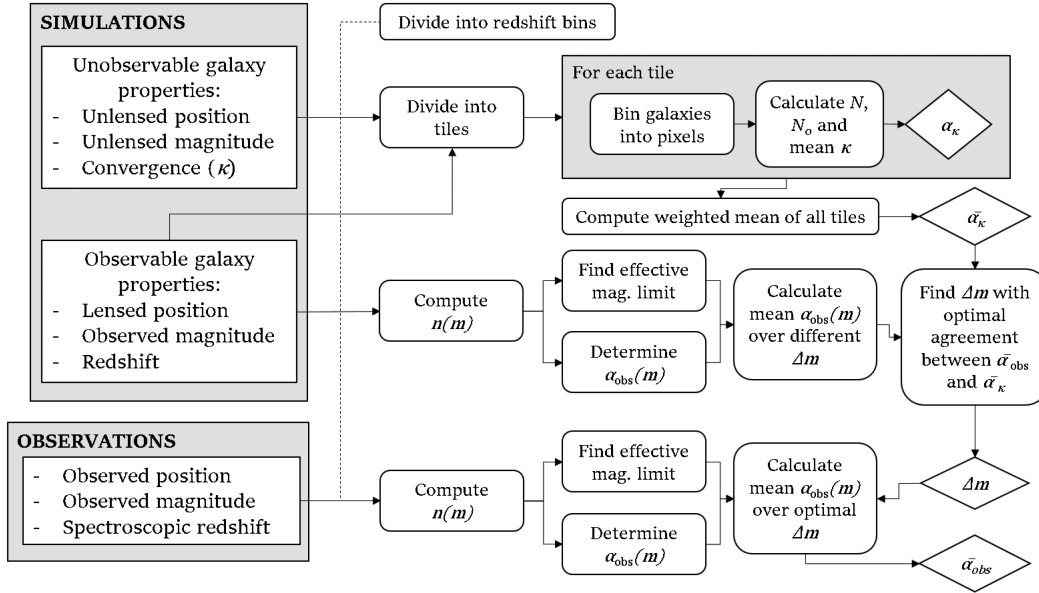
**Figure 2.** Normalized differential galaxy count distribution,  $n(i)$ , with respect to the  $i$ -band magnitude. The BOSS sample is shown in black, the MICE2 sample in red, while the flux-limited MICE2 sample is shown in yellow. In the top plot, we see the population of galaxies with  $0.2 < z \leq 0.5$  and at the bottom, the population of galaxies with  $0.5 < z \leq 0.75$ . The black cross indicates the effective magnitude limit determined for the BOSS sample by finding the faintest prominent peak in  $n(i)$ . The red triangle indicates the same for the MICE2 mock sample.

2020 for details). This allows us later to separate the effects of lensing dilution and magnification in the mock data. Unruh et al. 2020 recently showed that the weak lensing assumption ( $|\kappa| \ll 1$ ,  $|\gamma| \ll 1$ ) used to derive equation (4) and equation (12) might lead to biases when simulating magnified galaxy samples. Since 99.9 per cent of the galaxies in the MICE2 simulations have  $|\kappa| < 0.09$ , the assumption should still hold. However, it should be investigated in the future whether this is really the case.

Finally, we select two samples from this base catalogue, one with an arbitrary magnitude limit in the SDSS  $i$  band at  $m_i^{\text{mag}} \leq 20.2$  (applied in appendix A) and one that resembles the SDSS BOSS survey, using a target selection similar to Eisenstein et al. (2011; applied in Section 4). The details of this BOSS mock sample selection are summarized in van den Busch et al. (2020).

The  $i$ -band number counts of these two samples and the original BOSS data is shown in Fig. 2. In Fig. 2, it becomes apparent how the BOSS selection function differs from a flux-limited sample. The cut-off of the galaxy population at the magnitude limit is not as pronounced, while the  $n(i)$  no longer increases monotonically, especially at low redshifts. The galaxy counts per unit area as a function of redshift of the three samples is shown in Fig. 1. Here we see how with the BOSS selection function applied, the redshift distribution is altered in a highly non-linear manner causing it to be multi-peaked with a main peak at  $z \sim 0.5$ . The magnitude-limited sample, on the other hand, follows a roughly single-peaked distribution dominated by low redshift galaxies ( $z \sim 0.3$ ).





**Figure 3.** Flow chart outlining the method presented in this paper.  $N$  stands for the count of lensed galaxies,  $N_0$  refers to the counts of unlensed galaxies,  $\kappa$  to the convergence,  $\alpha_\kappa$  to the luminosity function slope determined from the known  $\kappa$ ,  $n(m)$  is the differential galaxy count distribution over magnitude,  $m$ ,  $\alpha_{\text{obs}}$  is the luminosity function slope as determined from  $n(m)$ .

Having knowledge of the underlying matter distribution allows us to compare estimates of the scale of flux magnification through  $\alpha_{\text{obs}}(m)$  from observables as given by equation (5) with the  $\alpha_\kappa$  estimate as given by equation (4). When analysing the MICE2 mock observations, we only consider the SDSS model  $i$ -band magnitude, due to a lack of available SDSS  $c$ model magnitudes from the simulations.

As a sanity check of our methods outlined in Section 3.3, we conduct an estimate of the magnification bias induced by a flux/magnitude-limited sample selection function on a galaxy survey over an eighth of the sky in appendix A. For this, we use the MICE2 simulations to obtain the positions and magnitudes of galaxies before and after magnification, while knowing the true underlying matter density. We set the magnitude limit in the  $i$  band to a magnitude of 20.2 (similar to the magnitude limit in the  $i$  band of the BOSS survey). We find that the calibrated  $\alpha_{\text{obs}}$  values accurately recover  $\alpha_\kappa$  near the faint limit. At the same time, the  $\alpha_{\text{obs}}$  estimates are robust over large changes in the calibration range chosen, showing that the power law approximation holds within  $\sim 1\sigma$  over  $\Delta i \sim 1$  and within  $\sim 2\sigma$  over the whole magnitude range for both  $z_{\text{low}}$  and  $z_{\text{high}}$ .

To conduct the analysis for the case where the target selection function is not flux or magnitude limited, we select a  $\sim 5000 \text{ deg}^2$  area from the MICE2 simulations and apply the aforementioned sample selection function to it. The  $i$  band magnitude distribution of the BOSS and MICE2 galaxies within each of the redshift bins is shown in Fig. 2. Here, we see that, although the overall shape of the population is similar between the BOSS and the MICE2 galaxies, the MICE2 objects are consistently shifted towards the fainter end of the distribution. This is at least partially caused by the fact that the BOSS magnitudes are  $i$ -band  $c$ model magnitudes and the MICE2 magnitudes are SDSS model  $i$ -band magnitudes. In addition, the MICE2 simulations with a BOSS-like selection function do not seem to capture the population of galaxies at the extremes of the magnitude distribution. Both of these biases might also be due to some assumptions in the galaxy formation and evolution models used in the MICE2 simulations. In addition, the fiducial cosmology assumed for

the simulations might not agree with the cosmological parameters preferred by the BOSS data. However, the method of calibrating the  $\alpha_{\text{obs}}$  estimates from the observations with the simulations is not sensitive to a constant shift in the distribution nor is it sensitive to the extremes of the magnitude distribution by construction.

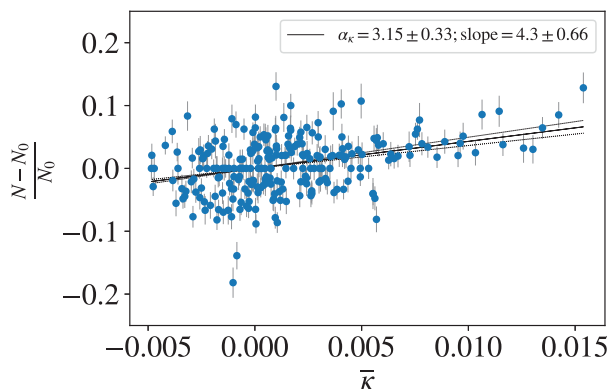
### 3.3 Calibration procedure on simulations

To calibrate the  $\alpha_{\text{obs}}$  obtained from observations, we first have to determine an accurate estimate of the underlying luminosity function slope,  $\alpha_\kappa$ , in the MICE2 simulations as given by equation (4). As outlined in Fig. 3, we first spatially bin the lensed and unlensed galaxy positions using HEALPIX at a resolution of  $\text{nside} = 64$  (Gorski et al. 2005). Within each bin/pixel, we evaluate lensed and unlensed cumulative galaxy counts,  $N$  and  $N_0$  respectively, as well as the average convergence,  $\kappa$ . We then perform a least squares linear fit of the relative difference between lensed and unlensed galaxy counts over the convergence,  $\kappa$ , to estimate  $\alpha_\kappa$  (as shown in Fig. 4). This is a consequence of the linearity between these two quantities which emerges in the weak lensing limit as given by equation (4).

In order to obtain better estimates of the uncertainties of  $\alpha$ , the HEALPIX pixels are grouped into tiles (HEALPIX pixels with a resolution of  $\text{nside} = 4$ ) for which we repeat the analysis independently each time. The weighted mean of these values obtained from each tile gives the final estimate for  $\alpha_\kappa$ ,  $\bar{\alpha}_\kappa$ , while the standard deviation between these values is used to estimate the uncertainty as given by

$$\sigma_\alpha^2 \equiv \frac{\sum_{i=1}^M \frac{(\alpha_i - \bar{\alpha})^2}{\sigma_i^2}}{(M-1) \sum_{i=1}^M \frac{1}{\sigma_i^2}}, \quad (13)$$

where  $\alpha_i$  are the  $\alpha$  estimates from each tile or bin,  $\sigma_i$  is their associated uncertainty,  $\bar{\alpha}$  is the weighted mean of the  $\alpha$  estimates and  $M$  is the number of tiles over which the analysis is repeated. When  $\sigma_i = \sigma$ , equation (13) reduces to the equation for the error of the mean, i.e.  $\sigma_\alpha \equiv \sigma_{\text{sd}}/\sqrt{M}$  where  $\sigma_{\text{sd}} \equiv (\sum_{i=1}^M (\alpha_i - \bar{\alpha})^2 / (M-1))^{1/2}$ .



**Figure 4.** Plot of the relative difference in galaxy counts per pixel over the mean convergence ( $\bar{\kappa}$ ) in each pixel for a HEALPIX pixelation with  $n_{\text{side}} = 64$  and in the zhigh redshift bin ( $0.5 < z \leq 0.75$ ). The graph only shows pixels within 1 of 28 tiles. The relative difference between the lensed and unlensed galaxy counts in each pixel are shown as blue points. The black line is fitted to the blue data points with equation (4) to give the  $\alpha_\kappa$  value shown in the legend.

As an alternative, one might think that it would be enough to assume that the uncertainty on the galaxy counts is given by a noise, which considers the correlation between the lensed and unlensed galaxy counts (which is shown in the errorbars of the data points in Fig. 4). We find, however, that this approach leads to underestimates of the uncertainties. Sampling  $\alpha_\kappa$  over many different areas in the sky gives a more conservative estimate of the uncertainty, while also accounting for the local fluctuations in the BOSS sample.

A possible cause for concern when comparing the magnified and unmagnified galaxy populations can be the edge cases where, for a given bin or pixel, the unmagnified galaxy number count  $N_0 = 0$ , while the magnified number counts  $N = 1$  or vice versa. These cases cause divergences in the relative difference and unrealistic uncertainties, since they introduce null denominators. For this reason, they are excluded in the analysis. In any case, the frequency of these occurrences is usually found to be negligible for the HEALPIX resolutions and redshift bins used in this work. Dividing the 5000  $\text{deg}^2$  MICE2 simulations into two redshift bins at a HEALPIX  $n_{\text{side}} = 64$ , there are none of these cases. While considering 19 redshift bins at the same HEALPIX resolution, only  $\sim 0.7$  per cent of the pixels have to be discarded.

### 3.4 Determining magnification bias from observations

After having determined the luminosity function slope,  $\bar{\alpha}_\kappa$ , from the simulations as described in Section 3.3, we estimate the optimal magnitude range,  $\Delta m$ , to calibrate the estimate of  $\alpha_{\text{obs}}$  from mock observations using  $\alpha_\kappa$ .

To do this, we first choose a magnitude band,  $m$ , that has been used to select (at least, partially) the galaxy sample of interest. Another magnitude band will carry less information about flux magnification. Then, we determine the discrete differential galaxy count distribution,  $n(m)$ , over the chosen magnitude,  $m$ , for a given redshift range. Subsequently, we find the magnitude at which the faintest most dominant peak in  $n(m)$  occurs. This value is considered to be the effective magnitude limit of the galaxy sample. From  $n(m)$ , we compute  $\alpha_{\text{obs}}(m)$  using equation (5). Thereafter, we calculate the weighted mean of  $\alpha_{\text{obs}}(m)$ ,  $\bar{\alpha}_{\text{obs}}$ , over all possible magnitude ranges,  $\Delta m$ , below the effective magnitude limit determined before.

In order to find the optimal  $\Delta m$  which will be used for the calibration of  $\bar{\alpha}_{\text{obs}}$  from the actual observations, we find the value of  $\bar{\alpha}_{\text{obs}}(\Delta m)$  which is in best statistical agreement with the value of  $\bar{\alpha}_\kappa$  determined previously for the same galaxy sample and redshift range. Therefore, the optimal  $\bar{\alpha}_{\text{obs}}(\Delta m)$  value is the one that minimizes the number of standard deviations it deviates from  $\bar{\alpha}_\kappa$ , i.e.  $|\bar{\alpha}_{\text{obs}}(\Delta m) - \bar{\alpha}_\kappa|/\sigma_{\alpha_\kappa}$ .

The reason behind choosing a magnitude range,  $\Delta m$ , relative to the effective magnitude limit of the differential galaxy count distribution,  $n(m)$ , for calibration is to account for one of the simplest forms of disagreement between the observed  $n(m)$  and the  $n(m)$  from mock observations. This disagreement is a constant shift in the domain of  $n(m)$ . For instance, such a shift exists between the  $n(m)$  from the BOSS and MICE2 samples that has been discussed in Section 3.2 and shown in Fig. 2 already. If we were to evaluate  $\bar{\alpha}_{\text{obs}}^{\text{MICE2}}$  and  $\bar{\alpha}_{\text{obs}}^{\text{BOSS}}$  over the same magnitude range, while disregarding the difference between their  $n(m)$  distributions, the  $\bar{\alpha}_{\text{obs}}$  estimates will be biased. This happens because we would be probing regimes of  $n(m)$  from the observed galaxy sample beyond or far below its magnitude limit when calculating  $\bar{\alpha}_{\text{obs}}$ . Other higher-order biases in the  $n(m)$  from mock observations may exist that would require more complex parametrizations of the calibration procedure. Nevertheless, in such cases, it might be more efficient and physically motivated to adjust the models used to produce the mock galaxy samples such that the agreement in  $n(m)$  improves up to a point where it can be mostly parametrized by a constant shift in the magnitude.

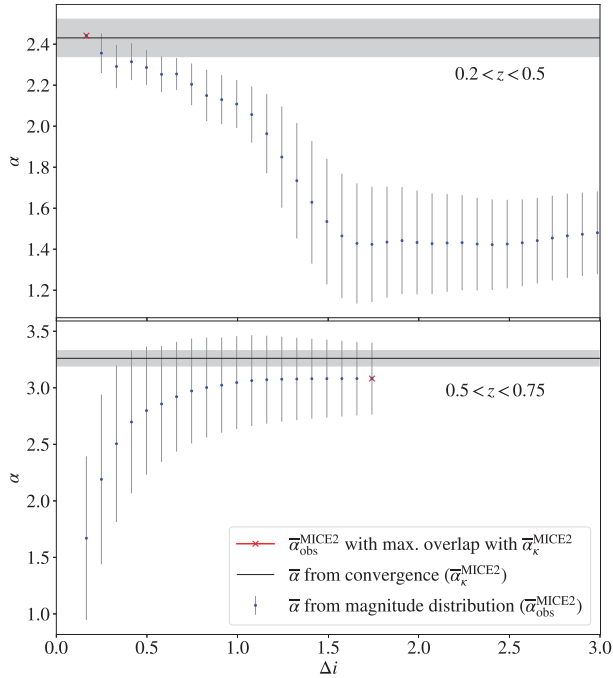
In any case, once the optimal  $\Delta m$  to reconcile  $\bar{\alpha}_{\text{obs}}$  and  $\bar{\alpha}_\kappa$  from the mocks has been determined, it may be used to calibrate  $\bar{\alpha}_{\text{obs}}$  from the observations. As summarized in the lower third of Fig. 3, we first compute  $n(m)$  for the given redshift range. We again find the faintest most dominant peak in  $n(m)$  and set it as the effective magnitude limit and evaluate  $\alpha_{\text{obs}}(m)$  from  $n(m)$ . Lastly, we calculate the weighted mean of  $\alpha_{\text{obs}}(m)$  over the optimal magnitude range below the effective magnitude limit,  $\Delta m$ , determined before from the simulations over the same redshift range. Thus, we produce the final  $\bar{\alpha}_{\text{obs}}$  estimate for that sample.

## 4 APPLICATIONS TO BOSS LENSES

We proceed to apply the method described in Sections 3.3 and 3.4 to the BOSS lens galaxy sample introduced in Section 3.1. The magnitude bands selected for this are cmodel magnitudes, since they are better indicators of the overall flux emitted by a galaxy. The specific magnitude band chosen is based on which band was used to select the dominant population within a sample. In other words, when working with LOWZ-dominated galaxy samples ( $z < 0.36$ ), we use the  $r$  band and when working with CMASS-dominated samples ( $z > 0.36$ ), we use the  $i$  band (Eisenstein et al. 2011). To allow for accurate forecasting of the KiDS-1000 + BOSS analysis (Heymans et al. 2020), we choose the same convention for the redshift bins:  $0.2 < z \leq 0.5$  and  $0.5 < z \leq 0.75$ . Consequently, both bins are dominated by CMASS galaxies, so we opt to use  $i$ -band magnitudes for the analysis of both samples.

As demonstrated in appendix A, for the flux-limited case, we can accurately and robustly estimate the magnitude of the magnification bias by determining the effective luminosity function slope  $\alpha$  through the weighted mean of  $\alpha_{\text{obs}}$  near the magnitude limit. In this section, we discuss whether the same can be said when applying a complex sample selection function that does not have a clear flux/magnitude limit such as in the case of the BOSS survey.

First, we directly estimate  $\alpha_\kappa$  from the MICE2 simulations following the approach outlined in Section 3.3. An example of this

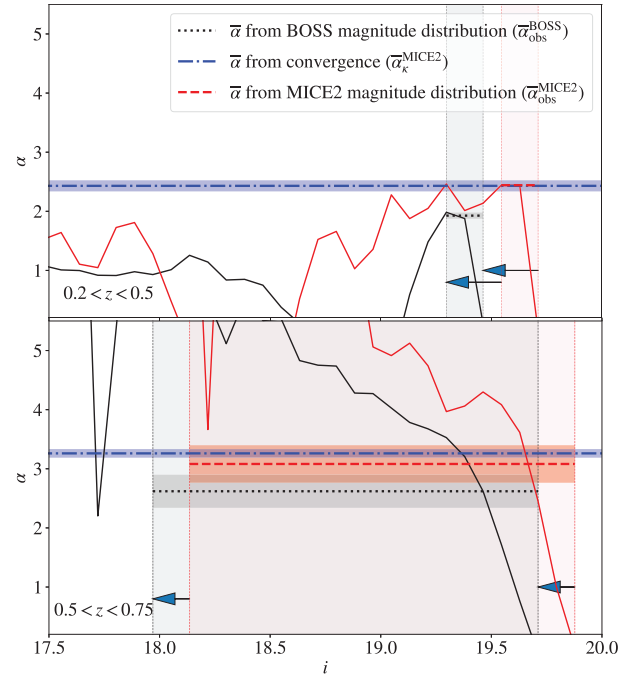


**Figure 5.**  $\bar{\alpha}_{\text{obs}}$  estimates from MICE2 simulations with the BOSS selection function over different  $i$ -band magnitude ranges below the turn-off magnitude,  $\Delta i$ , considered to calculate the weighted average. Two redshift ranges are considered: zlow with  $0.2 < z \leq 0.5$  (top) and zhigh with  $0.5 < z \leq 0.75$  (bottom). The red cross marks the  $\alpha$  estimate that overlaps the most with the  $\bar{\alpha}_k$  truth from the simulations (black vertical line).

is shown in Fig. 4, where we see the  $\alpha_k$  estimate within a single  $\sim 200 \text{ deg}^2$  tile containing 256 pixels within the zhigh bin. This procedure is repeated for each tile and redshift bin. Then, we find the weighted mean between the  $\alpha_k$  from each tile to determine the  $\bar{\alpha}_k$  for each redshift bin and its uncertainty given by equation (13). This gives  $\bar{\alpha}_k^{\text{zlow}} = 2.43 \pm 0.09$  and  $\bar{\alpha}_k^{\text{zhigh}} = 3.26 \pm 0.07$ .

Next, applying the procedure discussed in Section 3.4 and using the differential galaxy count distributions for each redshift bin shown in Fig. 2, we can estimate  $\bar{\alpha}_{\text{obs}}$ ; once for the simulated BOSS-MICE2 observations, and once for the actual BOSS observations. In Fig. 5, for zlow, we find that the estimate is optimal near the faint end of the count distribution, which is expected, since the assumed flux power law should be most accurate in the faint limit. However, this does not appear to be the case for the high redshift sample, zhigh. For this range, the estimate is optimal when considering the whole magnitude range up to the turn-off magnitude. This might be due to incompleteness in the sample and/or the complex selection, which flattens the observed number counts (Hildebrandt 2016).

Taking the magnitude range from the optimal  $\bar{\alpha}_{\text{obs}}^{\text{MICE2}}$  estimate to calibrate  $\bar{\alpha}_{\text{obs}}^{\text{BOSS}}$  gives the estimates shown in Fig. 6. For the MICE2 mocks, we find that  $\bar{\alpha}_k^{\text{zlow}} = 2.43 \pm 0.09$ , while  $\bar{\alpha}_{\text{obs}}^{\text{zlow}} = 2.442 \pm 0.002$ . In addition,  $\bar{\alpha}_k^{\text{zhigh}} = 3.26 \pm 0.07$ , while  $\bar{\alpha}_{\text{obs}}^{\text{zhigh}} = 3.08 \pm 0.32$  which indicates that the  $\alpha$  estimates obtained from observations using equation (5) are a good indicator of the scale of the magnification bias even when there is a complex sample selection function when they are properly calibrated. For this reason, we may consider the  $\bar{\alpha}_{\text{obs}}$  estimates given in Table 1 from the actual BOSS observations as unbiased indicators of the scale of the magnification bias. Note that the value for zlow, slightly deviates from the value of  $\bar{\alpha}_{\text{obs}}^{\text{BOSS}} = 1.80 \pm 0.15$  quoted in Joachimi et al. (2020), since there have been minor adjustments in the way peaks in  $n(m)$  are detected.



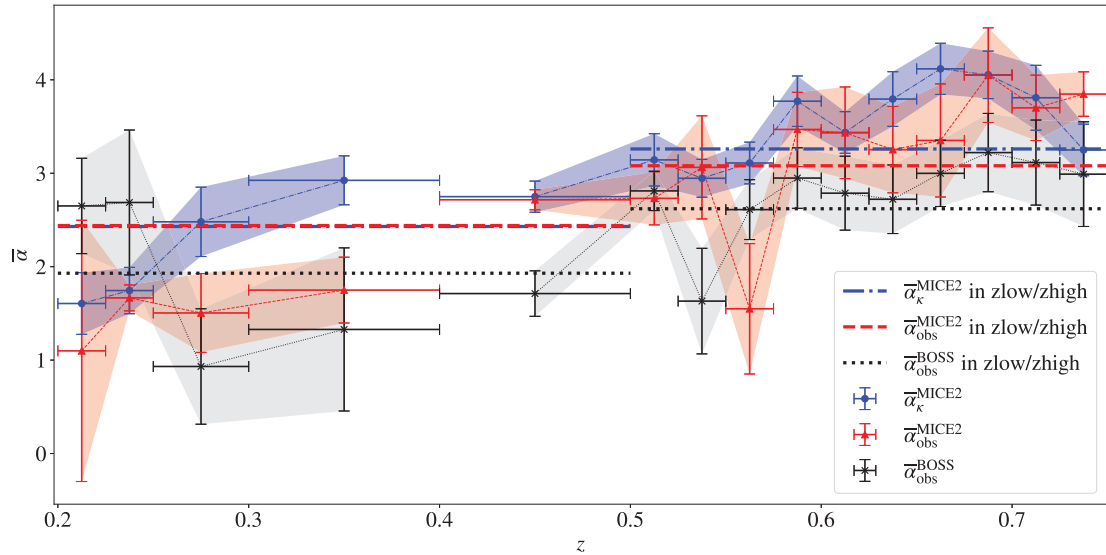
**Figure 6.** The slope of the BOSS luminosity function,  $\alpha$ , as a function of the  $i$ -band magnitude ( $i$ ) for two redshift bins:  $0.2 < z \leq 0.5$  (top) and  $0.5 < z \leq 0.75$  (bottom). The red line shows  $\alpha_{\text{obs}}(i)$  as given by equation (5) calculated from the MICE2 mocks, while the black line shown  $\alpha_{\text{obs}}(i)$  as determined from the BOSS DR12 photometric data. The grey vertical lines mark the upper and the lower bounds of the magnitude range used to find  $\bar{\alpha}_{\text{obs}}^{\text{BOSS}}$ , while the red vertical lines mark the upper and lower bounds of the highlighted magnitude range used to determine  $\bar{\alpha}_{\text{obs}}^{\text{MICE2}}$ . The arrows indicate the constant magnitude shift applied to reconcile the differential galaxy count distribution,  $n(m)$ , from observations with the  $n(m)$  from mocks. The dotted black horizontal line marks the  $\bar{\alpha}_{\text{obs}}$  estimate from BOSS galaxies, the dashed red horizontal line marks the  $\bar{\alpha}_{\text{obs}}$  estimate from MICE2 mock galaxies, and the blue dot-dashed horizontal line marks the effective  $\bar{\alpha}_k^{\text{MICE2}}$  determined from the weak lensing convergence with equation (4) and used to calibrate  $\bar{\alpha}_{\text{obs}}^{\text{MICE2}}$ .

**Table 1.** Table showing the effective luminosity function slopes derived from Fig. 6 for each redshift bin of the BOSS galaxy sample.

Bin	Redshift range	Luminosity function slope ( $\bar{\alpha}_{\text{obs}}^{\text{BOSS}}$ )
zlow	$0.2 < z \leq 0.5$	$1.93 \pm 0.05$
zhigh	$0.5 < z \leq 0.75$	$2.62 \pm 0.28$

This leads to a 16 per cent change in the amplitude of the mag. bias contribution, which has no effect on the KiDS-1000 analysis as the GGL contributions are marginal.

When comparing the  $\alpha_{\text{obs}}(i)$  curves for each bin in Fig. 6, one might notice that the turn-off near the effective magnitude limit is not as steep for zhigh as for zlow. This is due to the complex BOSS selection function that deviates particularly strongly from a simple flux limit at high redshifts. Here is where the semi-empirical calibration of the magnitude range considered in order to determine the effective luminosity function slope  $\bar{\alpha}_{\text{obs}}$  is especially relevant. As shown in Fig. 5, we find that for zhigh we get a more accurate  $\alpha$  estimate when considering the entire magnitude range  $\Delta i$  available below the effective magnitude limit that is in stark contrast with the results found for a flux-limited sample (see Fig. A2). The opposite is the case for zlow. As shown in Fig. 6, the double peak in the zlow bin



**Figure 7.** Plot of different  $\alpha$  estimates for 16 different redshift ( $z$ ) bins within  $0.2 < z \leq 0.75$ . The black crosses mark the  $\bar{\alpha}_{\text{obs}}^{\text{BOSS}}$  estimates from observations within each bin, the red triangles mark  $\bar{\alpha}_{\text{obs}}^{\text{MICE2}}$  estimates from mock observations, and the blue circles mark the true effective  $\bar{\alpha}_{\text{obs}}^{\text{MICE2}}$  determined from the weak lensing convergence with equation (4) and used to calibrate  $\bar{\alpha}_{\text{obs}}^{\text{MICE2}}$ . The values of  $\bar{\alpha}_{\text{obs}}^{\text{MICE2}}$  and  $\bar{\alpha}_{\text{obs}}^{\text{BOSS}}$  for the bins with  $z < 0.4$  have been derived from the differential galaxy count distribution with respect to the  $r$ -band magnitude,  $n(r)$ , while the values for the bins with  $z > 0.4$  have been derived from  $n(i)$ . The horizontal lines show the  $\alpha$  estimates from simulations obtained for the zlow bin ( $0.2 < z \leq 0.5$ ) and the zhigh bin ( $0.5 < z \leq 0.75$ ).

combined with a clearer ‘flux limit’ near the peak magnitude means that the power law model for the luminosity function holds best within a small magnitude range near the peak. In other words, the  $\Delta i$  intervals that provide the best agreement between  $\bar{\alpha}_{\text{obs}}$  and  $\bar{\alpha}_{\kappa}$  are also the magnitude intervals over which  $n(m)$  resembles a power law the most. Therefore, our method actively avoids basing its estimates on a magnitude domain where the power law approximation in equation (2) does not hold.

We note that in Fig. 2 the simulated and the observed differential count distributions do not quite match. The  $n(i)$  from MICE2 mock observations is shifted by a  $\Delta i \approx 0.2$  to the faint end with respect to the BOSS  $n(i)$ . This might be due to some limitations in the galaxy model of the MICE2 simulations. The fact that the  $n(m)$  from the mocks and observations do not match perfectly seems to be driving the discrepancy between  $\bar{\alpha}_{\text{obs}}^{\text{MICE2}}$  and  $\bar{\alpha}_{\text{obs}}^{\text{BOSS}}$  shown in Fig. 6. However, since our calibration is based on a magnitude range of a fixed width relative to the effective magnitude limit for each sample, the estimates are not sensitive to this apparent shift in the domain of  $n(i)$ . The only thing that can bias our estimates are any disagreements in higher-order derivatives of  $n(m)$  near the effective magnitude limit between observations and simulations. However, the uncertainties of  $\bar{\alpha}_{\text{obs}}$  from equation (13) are defined such that they consider the variations of  $\bar{\alpha}_{\text{obs}}$  within the calibration magnitude range.

In addition, to see how  $\alpha$  evolves over redshift within zlow and zhigh, we repeat this analysis of the BOSS sample again for a different choice of redshift bins producing the  $\alpha$  estimates shown in Fig. 7. Here, the edges of the 15 redshift bins are given by  $\{0.2, 0.225, 0.25, 0.3, 0.4, 0.5, 0.525, 0.55, 0.575, 0.6, 0.625, 0.65, 0.675, 0.7, 0.725, 0.75\}$ . Since the redshift bins between  $z = 0.2$  and  $z = 0.4$  are dominated by LOWZ galaxies, we choose a bin width of 0.1 instead of 0.025 between  $z = 0.3$  and  $z = 0.5$ . This is done to mitigate the sharp gradient changes in  $n(m)$  in the BOSS sample at redshifts near  $z = 0.36$ , i.e. at the boundary between the LOWZ and CMASS samples as shown in Fig. 1.

Fig. 7 shows how the effective luminosity function slope  $\bar{\alpha}_{\kappa}$  in the MICE2 sample varies smoothly. None the less,  $\bar{\alpha}_{\text{obs}}$  for MICE2 and

for BOSS varies more strongly with redshift, due to their sensitivity of small variations in  $n(m)$ . Also,  $\bar{\alpha}_{\kappa}^{\text{MICE2}}$  is consistent with  $\bar{\alpha}_{\text{obs}}^{\text{MICE2}}$  over most of the redshift range. However, for a few redshift bins,  $\bar{\alpha}_{\text{obs}}^{\text{MICE2}}$  is in a  $\sim 1\sigma$  to  $\sim 2\sigma$  tension with  $\bar{\alpha}_{\kappa}$  despite being calibrated to optimally overlap. Taking  $\bar{\alpha}_{\kappa}$  as the underlying truth, we consider  $\bar{\alpha}_{\text{obs}}^{\text{MICE2}}$  and  $\bar{\alpha}_{\text{obs}}^{\text{BOSS}}$  to be biased in these cases. This seems to be driven by small discrepancies between the faint-end of  $n(m)$  from MICE2 and the faint-end of  $n(m)$  from BOSS. These are then exacerbated, since a small change in the sample size can lead to radical changes in the gradient of the magnitude distribution  $n(m)$  of these galaxies, causing substantial biases in the  $\bar{\alpha}_{\text{obs}}$  estimates, as discussed in Hildebrandt (2016). None the less, these discrepancies become insignificant as we increase the sample size by widening the redshift bin width to the one used in the main analysis (i.e.  $0.2 < z \leq 0.5$  and  $0.5 < z \leq 0.75$ ). We also note that the  $\bar{\alpha}_{\text{obs}}^{\text{BOSS}}$  estimates for the  $0.2 < z \leq 0.225$  and  $0.225 < z \leq 0.25$  bins may be biased. This is the case, since the profile of  $n(m)$  as obtained from the MICE2 simulations deviates from the  $n(m)$  observed in BOSS more strongly than over the remaining redshift range. Hence, the calibration range determined through our method does not necessarily apply anymore (as already mentioned in Section 3.4) and the estimates may be inaccurate. To avoid this, we highlight the necessity for accurate cosmological simulations over the whole redshift domain.

## 5 MAGNIFICATION BIAS IN WEAK LENSING MEASUREMENTS

Having produced estimates for the effective luminosity function slope ( $\bar{\alpha}_{\text{obs}}$ ) for the BOSS DR12 galaxy sample, we now proceed to make forecasts of the importance of magnification bias in the GGL signals. The forecasts are produced from cross correlating source galaxies from weak lensing surveys with the BOSS lens samples considered in Section 4. First, we produce forecasts for the GGL signals for a KiDS-1000 + BOSS DR12 analysis as described in Joachimi et al. (2020). Secondly, we produce similar forecasts for a GGL analysis of HSC Wide + BOSS DR12 similar to Speagle et al. (2019), while using



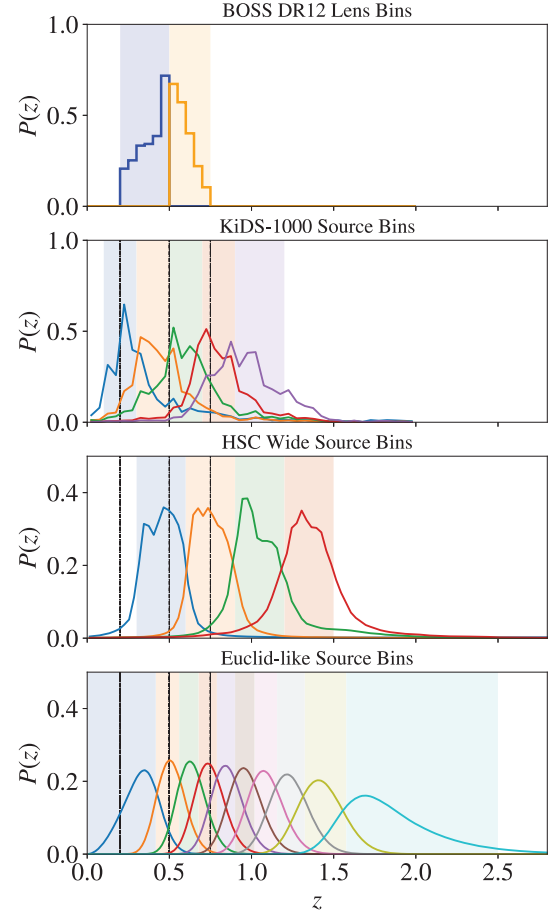
**Table 2.** Properties of the galaxy samples used to produce the galaxy–galaxy lensing forecasts.

Bin	$z$ range	$\bar{z}$	$z_{\text{med}}$	$n_{\text{gal}}$	$\sigma_{\epsilon, i}$
zlow	$0.2 < z_{\text{spec}} \leq 0.5$	0.38	0.37	0.014	-
zhigh	$0.5 < z_{\text{spec}} \leq 0.75$	0.60	0.55	0.016	-
KiDS1	$0.1 < z_{\text{phot}} \leq 0.3$	0.26	0.21	0.62	0.27
KiDS2	$0.3 < z_{\text{phot}} \leq 0.5$	0.40	0.36	1.18	0.26
KiDS3	$0.5 < z_{\text{phot}} \leq 0.7$	0.56	0.54	1.85	0.27
KiDS4	$0.7 < z_{\text{phot}} \leq 0.9$	0.79	0.75	1.26	0.25
KiDS5	$0.9 < z_{\text{phot}} \leq 1.2$	0.98	0.93	1.31	0.27
HSC1	$0.3 < z_{\text{phot}} \leq 0.6$	0.61	0.45	5.5	0.28
HSC2	$0.6 < z_{\text{phot}} \leq 0.9$	0.78	0.72	5.5	0.28
HSC3	$0.9 < z_{\text{phot}} \leq 1.2$	1.09	1.01	4.2	0.29
HSC4	$1.2 < z_{\text{phot}} \leq 1.5$	1.37	1.30	2.4	0.29
Euclid1	$0.001 < z_{\text{phot}} \leq 0.418$	0.33	0.21	3.0	0.21
Euclid2	$0.418 < z_{\text{phot}} \leq 0.560$	0.51	0.49	3.0	0.21
Euclid3	$0.560 < z_{\text{phot}} \leq 0.678$	0.63	0.62	3.0	0.21
Euclid4	$0.678 < z_{\text{phot}} \leq 0.789$	0.75	0.73	3.0	0.21
Euclid5	$0.789 < z_{\text{phot}} \leq 0.900$	0.85	0.84	3.0	0.21
Euclid6	$0.900 < z_{\text{phot}} \leq 1.019$	0.96	0.96	3.0	0.21
Euclid7	$1.019 < z_{\text{phot}} \leq 1.155$	1.09	1.09	3.0	0.21
Euclid8	$1.155 < z_{\text{phot}} \leq 1.324$	1.23	1.24	3.0	0.21
Euclid9	$1.324 < z_{\text{phot}} \leq 1.576$	1.42	1.45	3.0	0.21
Euclid10	$1.576 < z_{\text{phot}} \leq 2.500$	1.85	2.04	3.0	0.21

$\bar{z}$  stands for the mean redshift in each tomographic bin,  $z_{\text{med}}$  for the median redshift,  $n_{\text{gal}}$  for the galaxy number density in  $\text{arcmin}^{-2}$  following the definition from Heymans et al. (2012) and  $\sigma_{\epsilon, i}$  for the dispersion per ellipticity component. zlow and zhig are the lens bins based on the BOSS DR12 galaxy clustering data. The KiDS source bins have been defined in accordance with the methodology for the KiDS-1000 GGL analysis as given in Joachimi et al. (2020) and Heymans et al. (2020) based on the redshift calibration described in Hildebrandt et al. (2020) and Wright et al. (2020). The properties of the HSC source bins are based on the information provided in table 1 of the HSC Y1 cosmic shear analysis (Hikage et al. 2019) and the source  $P(z)$  distributions are based on the DEMp photometric redshifts. The tomographic bins for the Euclid forecasts are in accordance with the *Euclid* collaboration forecast choices (Blanchard et al. 2019). The *Euclid*  $P(z)$  distributions are determined using the fitting formula from Joachimi & Bridle (2010) assuming equi-populated binning with an overall median redshift of 0.8.

the source bins described in Hikage et al. (2019). Lastly, we produce GGL forecasts for a potential *Euclid*-like + DESI-like analysis using the galaxy sample properties defined in the *Euclid* collaboration forecast choices (Blanchard et al. 2019). The properties of all of the aforementioned galaxy samples are given in Table 2 and their redshift distributions,  $P(z)$ , are given in Fig. 8.

Throughout the forecasts, we assume a *Planck* 2018 TT,TE,EE + lowE flat  $\Lambda$ CDM cosmology (Planck Collaboration 2020) with  $\omega_b = 0.02236$ ,  $\omega_c = 0.1202$ ,  $h = 0.6727$ ,  $n_s = 0.9649$ ,  $\ln(10^{10} A_s) = 3.045$ ,  $\Omega_k = 0$ ,  $w = -1$ , and  $\sum m_\nu = 0.06 \text{ eV } c^{-2}$ . To model the cross-power spectrum between galaxy and matter distribution ( $P_{\text{gm}}$ ; e.g. Section 2.3), we split the power spectrum into linear and a non-linear part as outlined in Joachimi et al. (2020) based on Sánchez et al. (2017) and set  $b_1 = [2.1, 2.3]$ ,  $b_2 = [0.2, 0.5]$ , and  $\gamma_3 = [0.9, 0.1]$  where the first value of each vector corresponds to the first lens bin (zlow) and the second values to the second lens bin (zhig). These values follow the rounded best-fitting values from the cosmic shear and GGL analysis of KV450 + BOSS (Tröster et al. 2019). We use the halo and intrinsic alignment models described in Section 2 and set  $A_{\text{bary}} = 3.13$  (upper limit of the KiDS-1000 prior) and  $A_{\text{IA}} = 0.8$  (best estimate from Tröster et al. 2019).

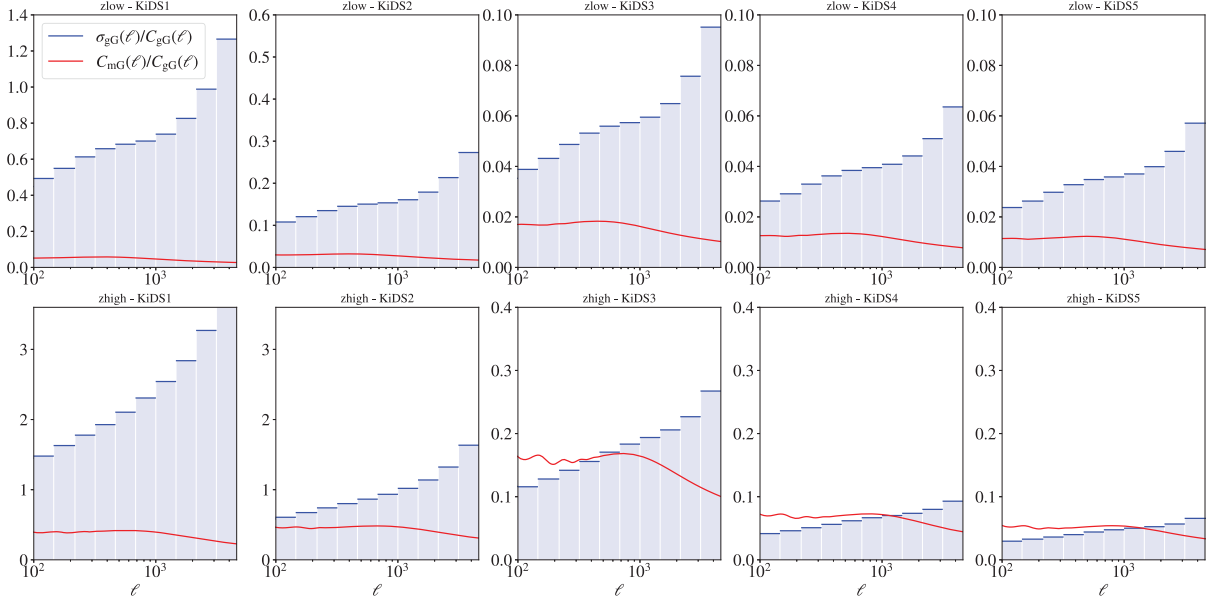


**Figure 8.** Redshift distributions  $P(z)$  for the lens and source galaxy samples used in the forecasts for the galaxy–galaxy lensing signal in a KiDS-1000+BOSS, HSC Wide + BOSS and *Euclid*-like + DESI-like analysis. The properties of these redshift distributions are given in Table 2. The black vertical dot–dashed lines show the limits of the BOSS lens bins for comparison with the source bins.

### 5.1 KiDS-1000 + BOSS forecasts

Following the approach outlined in Section 2.3, we propagate the  $\bar{\alpha}_{\text{obs}}$  measurements for zlow and zhig shown in Table 1 into angular power spectrum prediction for the galaxy–galaxy lensing signal. We then determine the ratio between the angular power spectrum correlating gravitational shear with the lensing-induced magnification bias in the lens sample,  $C_{\text{mG}}^{(ij)}(\ell)$ , and the angular power spectrum correlating the lens galaxy distribution and the source gravitational shear,  $C_{\text{gG}}^{(ij)}(\ell)$ , as shown in Fig. 9.

In order to put these contributions into perspective, we also estimate the statistical uncertainty in the GGL signal assuming shot and shape noise only (see for example Joachimi & Bridle 2010). We calculate this for six logarithmically spaced  $\ell$  bins per dex, while assuming the footprint area of the full KiDS survey,  $A = 1350 \text{ deg}^2$ . In Fig. 9, we then compare the relative magnification-shear signal to the relative GGL uncertainty for each  $\ell$  bin. The magnification-shear correlation found between these bins constitutes a few per cent contribution to the galaxy–galaxy lensing signal correlated with the zlow bin. To compare that to the shape and shot noise,  $\sigma_{\text{gG}}$ , we define the cumulative signal-to-noise ratio, SNR, within a range of angular



**Figure 9.** Magnification bias contribution  $C_{mG}(\ell)$  relative to the galaxy–galaxy lensing signal  $C_{gG}(\ell)$  over the angular scale  $\ell$  (in red) for the cross-correlations between the BOSS DR12 lens bins and the KiDS-1000 source bins assuming  $\bar{\alpha}_{obs}^{zlow} = 1.93$  and  $\bar{\alpha}_{obs}^{zhig} = 2.62$ . In blue, we show the expected relative uncertainty from shot and shape noise in the GGL signal,  $\sigma_{gG}(\ell)/C_{gG}(\ell)$ , within each  $\ell$  bin (six logarithmically spaced  $\ell$  bins per dex). The uncertainties are calculated for a KiDS footprint with an area of  $1350 \text{ deg}^2$ . The properties of the galaxy samples are given in Table 2.

scale,  $\ell_{min} < \ell < \ell_{max}$ , as follows:

$$\begin{aligned} \text{SNR}(\ell_{min} < \ell < \ell_{max}) &= \left( \frac{1}{K} \sum_i^K \text{SNR}_i^2 \right)^{1/2} \\ &= \left( \frac{1}{K} \sum_i^K \frac{C_{mG}^2(\ell_{min,i} < \ell < \ell_{max,i})}{\sigma_{gG}^2(\ell_{min,i} < \ell < \ell_{max,i})} \right)^{1/2}, \end{aligned} \quad (14)$$

where  $K$  is the number of  $\ell$  bins,  $i$  labels each  $\ell$  bin, and  $\ell_{min,i}$  and  $\ell_{max,i}$  mark the lower and upper limits of each bin, respectively. For the correlations with the zlow bin, this implies a cumulative signal-to-noise ratio for  $100 < \ell < 4600$  between 0.1 and 0.3. This contribution becomes larger for the high-redshift source bin (zhig), from  $\sim 5$  per cent to  $\sim 20$  per cent of the GGL signal, while the shot and shape noise is of a similar scale. Hence, the cumulative  $\text{SNR}(100 < \ell < 4600) = 0.2$  for the correlation between the zhig and the first KiDS redshift bin, while the cumulative  $\text{SNR}(100 < \ell < 4600) = 1.1$  between the zhig and the fifth KiDS bin. At the same time, these  $\alpha$  values lead to a maximal contribution of the magnification bias to the clustering signal of  $\sim 0.6$  per cent (Joachimi et al. 2020). Even though we are assuming the area of the full  $1350 \text{ deg}^2$  KiDS footprint, these contributions to the GGL signal by magnification are large enough to prompt the consideration through modelling in the analysis of this systematic in the KiDS-1000 + BOSS analysis outlined in Joachimi et al. (2020). None the less, since the analysis shown here already provides an accurate estimate for the magnitude of the magnification bias, the contribution to the GGL signal in each bin can simply be fixed and added to the overall GGL angular power spectrum without the need to add any more free parameters in the astrophysical models.

We note the oscillations at low  $\ell$  for some of the zhig correlations in Fig. 9. These originate from fluctuations from a power law of  $< 1$  per cent in  $C_{gG}(\ell)$  for  $100 < \ell < 500$ . They can be attributed to baryonic acoustic oscillations, BAO, as their amplitude decreases with  $\omega_b$ . In Fig. 9 and subsequent forecasts discussed Sections 5.2 and 5.3, the fluctuations in  $C_{mG}(\ell)/C_{gG}(\ell)$  at low  $\ell$  appear to be increased to amplitudes  $> 1$  per cent from the mean. This is caused

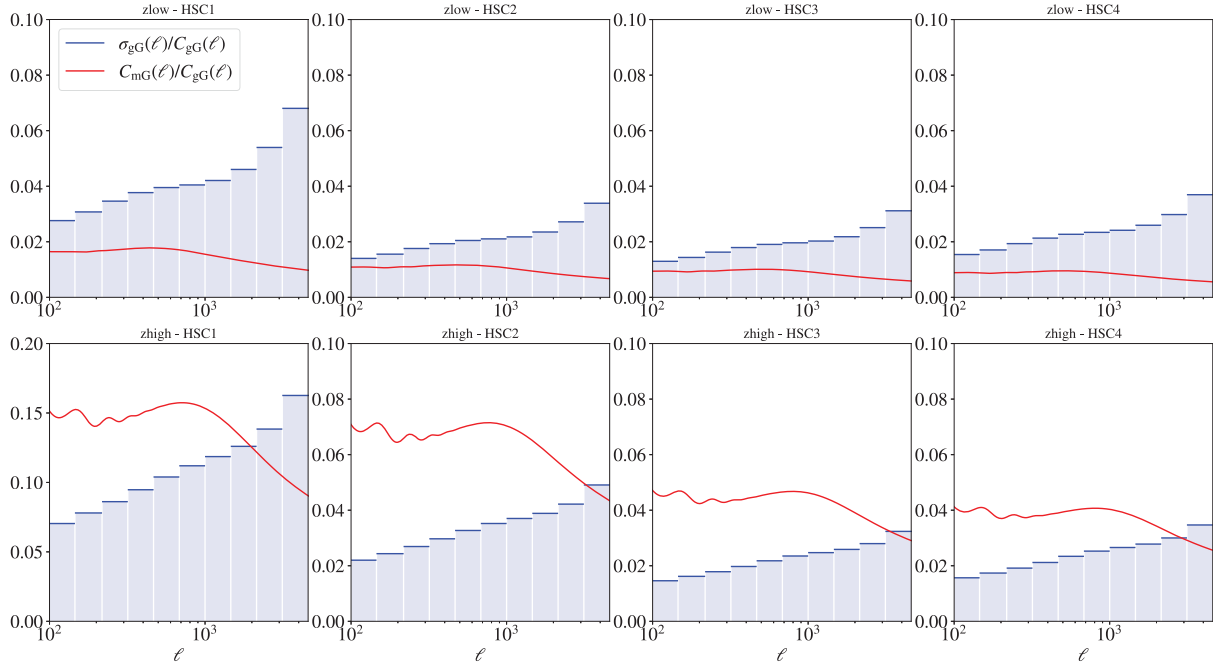
by the non-BAO signal in  $C_{gG}(\ell)$  being approximately proportional to  $C_{mG}(\ell)$  at low  $\ell$ . Hence, after taking their ratio, the only signal that does not approximately cancel is the BAO signal in  $C_{gG}(\ell)$ . In any case, the variations in  $C_{gG}(\ell)$  are well below the uncertainties over that range (which are typically  $\gg 2$  per cent), so they would be undetectable for now.

## 5.2 HSC Wide + BOSS forecasts

We repeat the analysis for Section 5.1, considering the HSC Wide source bins. Fig. 10 shows the ratio between  $C_{mG}^{(i,j)}(\ell)$  and  $C_{gG}^{(i,j)}(\ell)$  together with the relative uncertainty in the GGL signal for each  $\ell$  bin assuming a full footprint area of  $1400 \text{ deg}^2$  (Aihara et al. 2018) as well as the galaxy sample properties shown in Table 2. Similar to KiDS, we find that the magnification-shear signal only contributes about  $\sim 2$  per cent to the GGL signal correlated with the zlow lens bin (giving a cumulative SNR within  $100 < \ell < 4600$  between 0.4 and 0.5). In correlations with the zhig lens bin, the contribution of the magnification-shear signal is larger and between  $\sim 5$  per cent and  $\sim 20$  per cent which is considerable above the shape and shot noise (with a cumulative SNR within  $100 < \ell < 4600$  between 1.3 and 2.0). It is significant enough to give grounds for the consideration of this systematic during future GGL analyses that cross-correlate the HSC Wide sample with the BOSS DR12 or a similarly selected lens sample.

## 5.3 Euclid-like survey + DESI-like survey forecasts

We produce forecasts for a GGL analysis with Stage-IV (Albrecht et al. 2006), assuming lens and source samples akin to DESI (Aghamousa et al. 2016) and Euclid (Laureijs et al. 2011), respectively. We repeat the analysis shown in Sections 5.1 and 5.2 for the Euclid-like source bins described in Table 2 and in Fig. 8. We consider a footprint overlap between our source and lens sample of  $6000 \text{ deg}^2$ , which is roughly the expected overlap between Euclid and DESI (Levi et al. 2013; Aghamousa et al. 2016). Therefore,



**Figure 10.** Same as Fig. 9, but for HSC source bins as defined in Table 2 and Fig. 8.

the fictitious BOSS/DESI-like galaxy sample we are considering here has all the properties of the BOSS lens sample, but has the planned DESI footprint. Although DESI will probe higher redshifts and fainter galaxies than BOSS, it will be similar to BOSS in that it will not be a purely flux-limited survey. Targets in DESI are selected using a combination of different band magnitudes depending on the galaxy type and redshift range that is being observed (for more details see Aghamousa et al. 2016). For this reason, the magnification bias in the DESI sample cannot be modelled analytically either, warranting an analysis similar to the one discussed here. The *Euclid*-like source sample used in this work is designed to be split into the same redshift bins as suggested by *Euclid* collaboration forecast choices (Blanchard et al. 2019). In addition, within each bin, the median redshift is chosen to be in agreement with the one expected for the *Euclid* sources.

Considering six logarithmically spaced  $\ell$  bins per dex (as in the previous sections), we obtain the magnification-shear signal forecasts shown in Fig. 11. We see that the magnification-shear signal constitutes a considerable systematic when correlating with the zlow bin, since the observed cumulative SNR on scales within  $100 < \ell < 4600$  is between 0.3 and 0.7. The magnification bias signal becomes strong enough for correlations with zhigh, it would be a detectable signal (with the cumulative SNR within  $100 < \ell < 4600$  ranging from 1.5 when correlating zhigh and Euclid1 to 2.8 when correlating zhigh and Euclid10). This might require any future GGL analysis of *Euclid* + BOSS or *Euclid* + DESI data to allow for the  $\alpha$  parameters to freely vary as a nuisance parameter in order to properly account for this systematic. The method outlined in this paper could be used to set informative priors on the  $\alpha$  values within each lens bin.

## 6 CONCLUSIONS

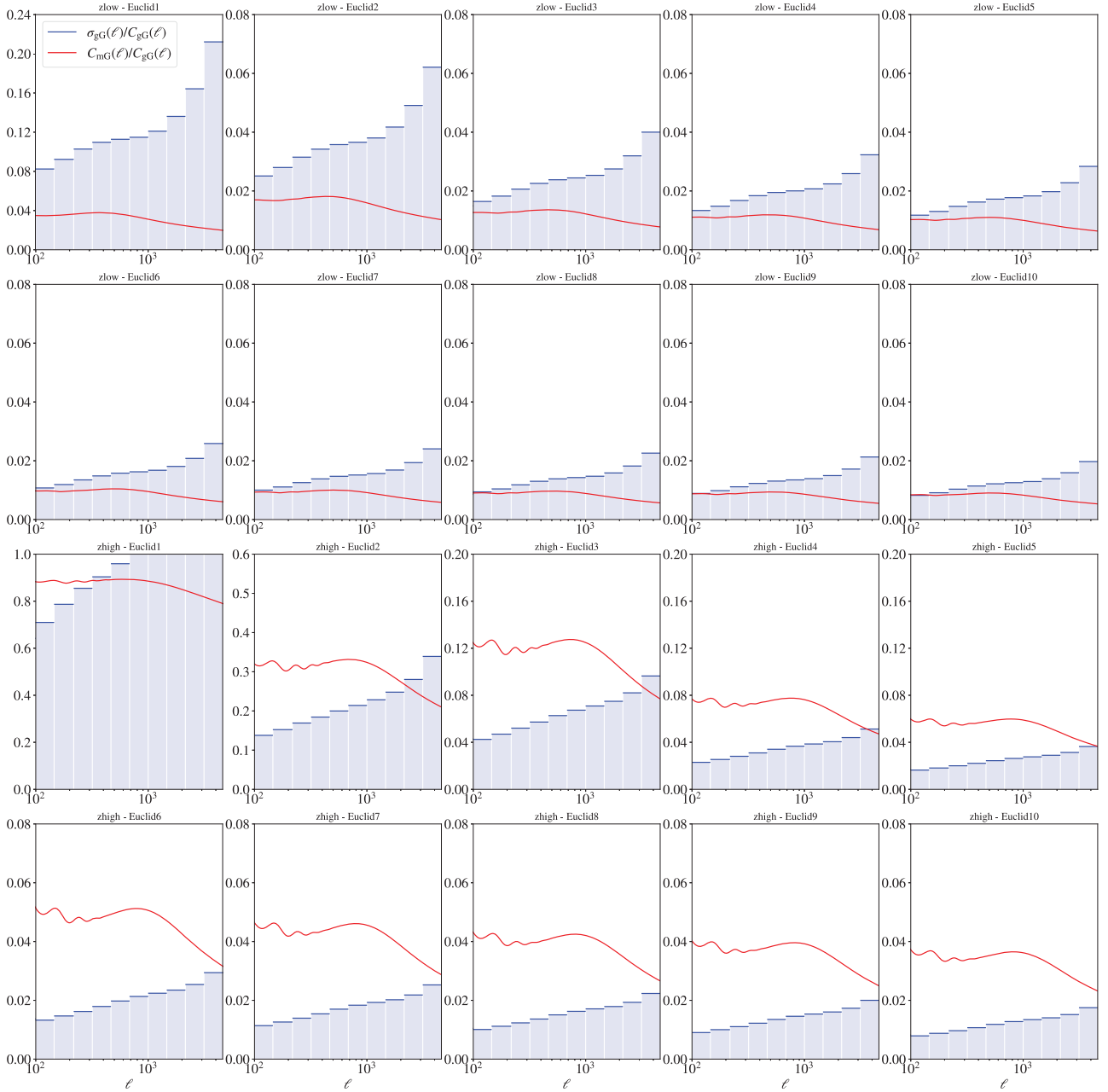
In this paper, we have introduced a novel method to estimate the effective luminosity function slope,  $\alpha$ , of galaxy samples that have been defined with a complex selection function that is not simply

flux/magnitude-limited. The method calibrates the  $\alpha$  estimates from observables with accurate cosmological simulations with the same sample selection. This expands upon previous work where the flux magnification was only measured for flux-limited cases or found to be inaccurate in non-flux-limited cases (Hildebrandt 2016).

The new method determines the underlying slope of the luminosity function of the simulated galaxy sample ( $\alpha_\kappa$ ) from unobservable properties such as the convergence,  $\kappa$ , and the unlensed galaxy position. It then finds the magnitude range relative to the magnitude limit over which the resulting  $\alpha_{\text{obs}}$  as calculated from the observable differential galaxy count distribution,  $n(m)$ , best agrees with  $\alpha_\kappa$ . Finally, the same relative magnitude range is used to determine  $\alpha_{\text{obs}}$  from the observed galaxy sample.

A few things should be considered when employing this method. We find that the magnitude ranges up to the effective magnitude limit that are determined to be optimal from the simulations in order to calibrate  $\alpha_{\text{obs}}$  are only valid for a given redshift range, a given sample selection function, and a given galaxy sample for which weak lensing simulations are available. Thus, it is important to note that this method cannot be generalized trivially, as it requires the availability of accurate cosmological simulations to assure consistency between the two independent  $\alpha$  estimates,  $\bar{\alpha}_{\text{obs}}$  and  $\bar{\alpha}_\kappa$ . None the less, when simulations are available, it provides a robust estimate of the scale of the magnification bias for non-flux-limited surveys such as BOSS.

Applying our calibration method to the BOSS DR12 sample split into two redshift bins, we find that  $\bar{\alpha}_{\text{obs}} = 1.93 \pm 0.05$  for  $0.2 < z \leq 0.5$  and  $\bar{\alpha}_{\text{obs}} = 2.62 \pm 0.28$  for  $0.5 < z \leq 0.75$  leading to a contribution to the galaxy–galaxy lensing signal of up to  $\sim 2$  per cent for KiDS-1000 and HSC Wide sources correlated with the  $0.2 < z \leq 0.5$  lens bin. Although the contribution can go up to  $\sim 20$  per cent when correlating KiDS-1000 and HSC Wide sources with the  $0.5 < z \leq 0.75$  BOSS lens bin, the magnification-shear signal can go above the noise with a cumulative SNR going up to 1.1 and 2.0 for KiDS-1000 and HSC Wide, respectively. Hence, both for KiDS-1000 and HSC Wide, the magnification-shear signal appears to be dominant enough to warrant the modelling of this systematic in future GGL



**Figure 11.** Same as Fig. 9, but for *Euclid*-like source bins as defined in Table 2 and Fig. 8.

analyses involving BOSS lenses, as was already done in the recent KiDS-1000 analysis (Heymans et al. 2020; Joachimi et al. 2020). This necessity becomes even more evident in the forecasts for a GGL analysis of *Euclid*-like sources with DESI-like lenses. In this case, the magnification-shear signal is either a considerable systematic when correlating with the zlow bin (with a cumulative SNR of around 0.5), or it even becomes a detectable signal when correlated the source bins with zhig giving cumulative SNRs around 2 that can go up to 2.8. This might require any future GGL analysis incorporating *Euclid* and any highly selected lens sample (e.g. DESI or BOSS) to allow for the effective luminosity function slope ( $\alpha$ ) of each lens sample to vary freely within the model using informative priors based on an analysis similar to the one conducted in this paper. These results are in line with Duncan et al. (2014) as well as the recent findings from

Mahony et al. (in prep.) where it was determined that the inclusion of the magnification bias in the modelling for surveys such as the next generation of surveys is necessary to accurately infer cosmological parameters.

We expect similar conclusions for other surveys. It might be desirable to estimate the magnification bias using the methodology outlined in this paper in clustering and GGL analyses based on DES REDMAGIC lens galaxies such as the ones described in Clampitt et al. (2017), Elvin-Poole et al. (2018), and Prat et al. (2018), since it also follows a complex selection function (Roza et al. 2016). The SNR should be comparable to HSC and KiDS, so the magnification bias will not have to be included as a free parameter. On the other hand, for surveys such as LSST (Ivezić et al. 2008; Abell et al. 2009) and the Nancy Grace Roman Space Telescope (formerly known as



WFIRST; Spergel et al. 2015), it may become necessary to make the  $\alpha$  of the lens galaxy samples a nuisance parameter in any clustering or GGL analysis, as we suggest for a *Euclid* + DESI-like analysis.

## ACKNOWLEDGEMENTS

We thank our referee for a constructive and comprehensive report. We would also like to thank Chieh-An Lin for helpful discussions. MWK thanks the Science and Technology Facilities Council for support in the form of a PhD Studentship. We acknowledge support from the European Research Council under grant numbers 647112 (CH, MA, TT) and 770935 (HH, AW, JvdB), as well as the Deutsche Forschungsgemeinschaft (HH, Heisenberg grant Hi 1495/5-1). CH and SU also acknowledge support from the Max Planck Society and the Alexander von Humboldt Foundation in the framework of the Max Planck-Humboldt Research Award endowed by the Federal Ministry of Education and Research.

The MICE simulations have been developed at the MareNostrum supercomputer (BSC-CNS) thanks to grants AECT-2006-2-0011 through AECT-2015-1-0013. This work has made use of CosmoHub (Carretero et al. 2017). CosmoHub has been developed by the Port d'Informació Científica (PIC), maintained through a collaboration of the Institut de Física d'Altes Energies (IFAE) and the Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT), and was partially funded by the 'Plan Estatal de Investigación Científica y Técnica y de Innovación' program of the Spanish government.

## DATA AVAILABILITY

The SDSS-III BOSS DR12 data (Eisenstein et al. 2011; Dawson et al. 2012) underlying this article is available at <https://www.sdss.org/dr12/> and the data sets generated by the MICE2 simulations (Carretero et al. 2015; Crocce et al. 2015; Fosalba et al. 2015a, b; Hoffmann et al. 2015) used in this work are accessible on CosmoHub (Carretero et al. 2017; <https://cosmohub.pic.es/home>). The code used for the analysis presented in this paper can be found in the MAGBET GitHub repository (<https://github.com/mwiet/MAGBET>).

## REFERENCES

Abbott T. et al., 2018, *Phys. Rev. D*, 98, 043526  
 Abbott T. et al., 2019a, *Phys. Rev. D*, 100, 023541  
 Abbott T. et al., 2019b, *Phys. Rev. Lett.*, 122, 171301  
 Abell P. A. et al., 2009, preprint (arXiv:0912.0201)  
 Aghamousa A. et al., 2016, preprint (arXiv:1611.00036)  
 Aihara H. et al., 2018, *PASJ*, 70, S8  
 Alam S. et al., 2015, *ApJS*, 219, 12  
 Alam S. et al., 2017, *MNRAS*, 470, 2617  
 Albrecht A. et al., 2006, preprint (astro-ph/0609591)  
 Asgari M., et al., 634, 2020, *A&A*, A127  
 Bartelmann M., Schneider P., 2001, *Phys. Rep.*, 340, 291  
 Beutler F. et al., 2017, *MNRAS*, 466, 2242  
 Binggeli B., Sandage A., Tammann G., 1988, *ARA&A*, 26, 509  
 Blanchard A. et al., 2019, preprint (arXiv:1910.09273)  
 Bridle S., King L., 2007, *New J. Phys.*, 9, 444  
 Broadhurst T., Lehar J., 1995, *ApJ*, 450, L41  
 Brown M., Taylor A., Hambly N., Dye S., 2002, *MNRAS*, 333, 501  
 Carretero J., Castander F. J., Gaztañaga E., Crocce M., Fosalba P., 2015, *MNRAS*, 447, 646  
 Carretero J., et al., 2017, PoS, EPS-HEP2017, 488  
 Chiu I. et al., 2016, *MNRAS*, 457, 3050  
 Clampitt J. et al., 2017, *MNRAS*, 465, 4204

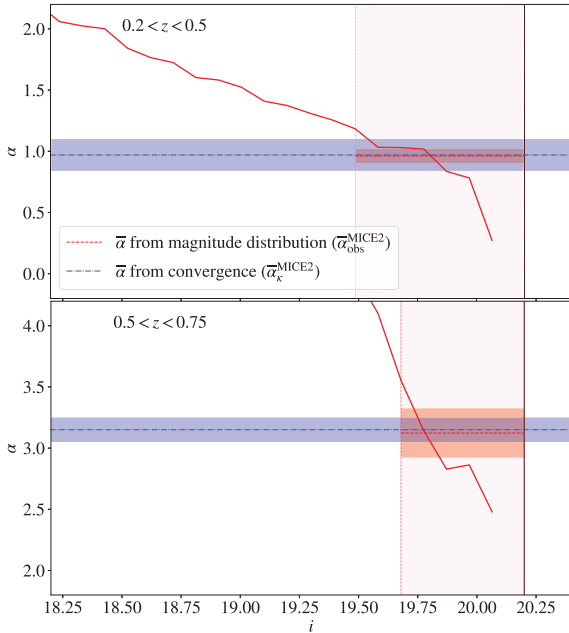
Crocce M., Castander F. J., Gaztañaga E., Fosalba P., Carretero J., 2015, *MNRAS*, 453, 1513  
 Dawson K. S. et al., 2012, *AJ*, 145, 10  
 Duncan C. A., Joachimi B., Heavens A. F., Heymans C., Hildebrandt H., 2014, *MNRAS*, 437, 2471  
 Eisenstein D. J. et al., 2011, *AJ*, 142, 72  
 Elvin-Poole J. et al., 2018, *Phys. Rev. D*, 98, 042006  
 Flaugher B. et al., 2015, *AJ*, 150, 150  
 Fosalba P., Crocce M., Gaztañaga E., Castander F. J., 2015a, *MNRAS*, 448, 2987  
 Fosalba P., Gaztañaga E., Castander F. J., Crocce M., 2015b, *MNRAS*, 447, 1319  
 Garcia-Fernandez M. et al., 2018, *MNRAS*, 476, 1071  
 Gorski K. M., Hivon E., Banday A. J., Wandelt B. D., Hansen F. K., Reinecke M., Bartelmann M., 2005, *ApJ*, 622, 759  
 Hambly N. C. et al., 2001, *MNRAS*, 326, 1279  
 Heymans C. et al., 2012, *MNRAS*, 427, 146  
 Heymans C. et al., 2021, *A&A*, 646, A140  
 Hikage C. et al., 2019, *PASJ*, 71, 43  
 Hildebrandt H., 2016, *MNRAS*, 455, 3943  
 Hildebrandt H., Van Waerbeke L., Erben T., 2009, *A&A*, 507, 683  
 Hildebrandt H. et al., 2017, *MNRAS*, 465, 1454  
 Hildebrandt H. et al., 2021, *A&A*, 647, A124  
 Hirata C. M., Seljak U., 2004, *Phys. Rev. D*, 70, 063526  
 Hoffmann K., Bel J., Gaztañaga E., Crocce M., Fosalba P., Castander F. J., 2015, *MNRAS*, 447, 1724  
 Howlett C., Lewis A., Hall A., Challinor A., 2012, *J. Cosmol. Astropart. Phys.*, 2012, 027  
 Huff E. M., Graves G. J., 2013, *ApJ*, 780, L16  
 Hui L., Gaztañaga E., LoVerde M., 2007, *Phys. Rev. D*, 76, 103502  
 Ivezić v. et al., 2008, preprint (arXiv:0805.2366v4)  
 Joachimi B., Bridle S., 2010, *A&A*, 523, A1  
 Joachimi B., Mandelbaum R., Abdalla F., Bridle S., 2011, *A&A*, 527, A26  
 Joachimi B. et al., 2021, *A&A*, 646, A129  
 Kaiser N., 1992, *ApJ*, 388, 272  
 Kiessling A. et al., 2015, *Space Sci. Rev.*, 193, 67  
 Kirk D. et al., 2015, *Space Sci. Rev.*, 193, 139  
 Kuijken K. et al., 2015, *MNRAS*, 454, 3500  
 Laureijs R. et al., 2011, preprint (arXiv:1110.3193)  
 Levi M. et al., 2013, preprint (arXiv:1308.0847)  
 Lewis A., Bridle S., 2002, *Phys. Rev. D*, 66, 103511  
 Lewis A., Challinor A., Lasenby A., 2000, *ApJ*, 538, 473  
 LoVerde M., Afshordi N., 2008, *Phys. Rev. D*, 78, 123506  
 Mead A., Heymans C., Lombriser L., Peacock J., Steele O., Winther H., 2016, *MNRAS*, 459, 1468  
 Mead A., Peacock J., Heymans C., Joudaki S., Heavens A., 2015, *MNRAS*, 454, 1958  
 Mahony C. et al. in prep.  
 Planck Collaboration, 2020, *A&A*, 641, A6  
 Prat J. et al., 2018, *Phys. Rev. D*, 98, 042005  
 Reid B. et al., 2016, *MNRAS*, 455, 1553  
 Rozo E. et al., 2016, *MNRAS*, 461, 1431  
 Sánchez A. G. et al., 2017, *MNRAS*, 464, 1640  
 Schechter P., 1976, *ApJ*, 203, 297  
 Schmidt F., Leauthaud A., Massey R., Rhodes J., George M. R., Koekemoer A. M., Finoguenov A., Tanaka M., 2011, *ApJ*, 744, L22  
 Scranton R. et al., 2005, *ApJ*, 633, 589  
 Speagle J. S. et al., 2019, *MNRAS*, 490, 5658  
 Spergel D. et al., 2015, preprint (arXiv:1503.03757)  
 Stoughton C. et al., 2002, *AJ*, 123, 485  
 Thiele L., Duncan C. A., Alonso D., 2020, *MNRAS*, 491, 1746  
 Tröster T. et al., 2020, *A&A*, 633, L10  
 Tröster T. et al., 2020, preprint (arXiv:2010.16416v2)  
 Troxel M., Ishak M., 2015, *Phys. Rep.*, 558, 1  
 Unruh S., Schneider P., Hilbert S., Simon P., Martin S., Puertas J. C., 2020, *A&A*, 638, A96  
 Vakili M. et al., 2020, preprint (arXiv:2008.13154)  
 van den Busch J. L. et al., 2020, *A&A*, 642, A200

Wright A. H., Hildebrandt H., van den Busch J. L., Heymans C., 2020, *A&A*, 637, A100  
 Ziour R., Hui L., 2008, *Phys. Rev. D*, 78, 123517

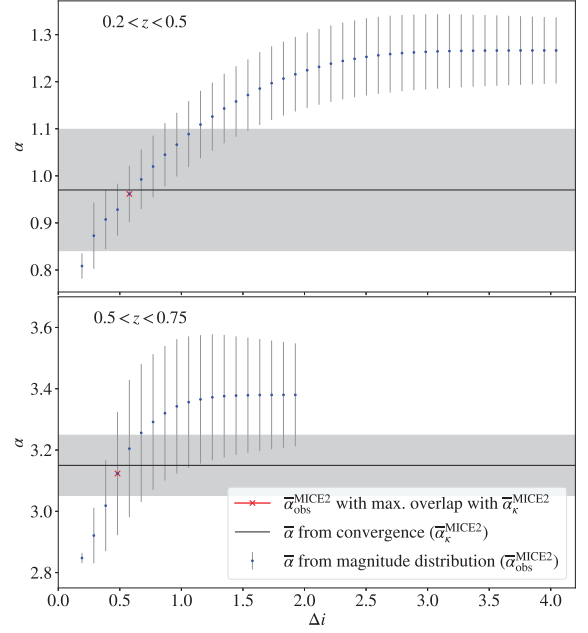
## APPENDIX A: FLUX-LIMITED CASE

As discussed in Section 3.2, we conduct a sanity check of our method by comparing the effective  $\bar{\alpha}_\kappa$  to  $\bar{\alpha}_{\text{obs}}$  for each redshift bin given a simulated magnitude-limited ( $i < 20.2$ ) galaxy population spanning the whole sky. This sample is also based on MICE2 simulations. We estimate  $\bar{\alpha}_\kappa$  from the known matter convergence  $\kappa$  and the relative difference between the lensed and unlensed cumulative galaxy number counts finding that  $\bar{\alpha}_\kappa^{\text{zlow}} = 0.97 \pm 0.13$  in the zlow bin ( $0.2 < z \leq 0.5$ ) and that  $\bar{\alpha}_\kappa^{\text{zhigh}} = 3.15 \pm 0.10$  in the zhight bin ( $0.5 < z \leq 0.75$ ).

The  $\bar{\alpha}_\kappa$  values are compared to  $\bar{\alpha}_{\text{obs}}$  in Fig. A1. We find that in zlow,  $\bar{\alpha}_{\text{obs}}$  optimally overlaps with the  $\bar{\alpha}_\kappa$  estimate from the convergence when taking the weighted mean of  $\alpha_{\text{obs}}(m)$  over a magnitude range of  $\Delta i = 0.67$  below the effective magnitude limit; giving  $\bar{\alpha}_{\text{obs}}^{\text{zlow}} = 0.96 \pm 0.06$ . This agrees well with the  $\bar{\alpha}_\kappa$  of galaxies in zlow ( $\bar{\alpha}_\kappa^{\text{zlow}} = 0.97 \pm 0.13$ ). The agreement is similarly good in the zhight bin where the optimal  $\bar{\alpha}_{\text{obs}}$  is computed over a  $\Delta i = 0.48$  and found to be  $\bar{\alpha}_{\text{obs}}^{\text{zhight}} = 3.12 \pm 0.20$ . The excellent agreement between  $\bar{\alpha}_\kappa$  and when evaluating near the faint end of the sample ( $\Delta i < 0.7$ ) reinforces that equation (4) and (5) indeed describe the same  $\alpha$ ; at least for the flux-limited. Such a good agreement is not really surprising, since the underlying assumptions leading to equation (4) ( $|\kappa| \ll 1$ ;  $|\gamma| \ll 1$ ) are ingrained in the way the MICE2 simulations determine the magnified magnitude and position of galaxy (Fosalba et al. 2015a). None the



**Figure A1.** The slope of the luminosity function,  $\alpha$ , as a function of the  $i$ -band magnitude,  $i$ , for the magnitude-limited case. Two redshift bins are considered:  $0.2 < z \leq 0.5$  (top) and  $0.5 < z \leq 0.75$  (bottom). The vertical black line marks the magnitude limit at  $i = 20.2$  and the dashed red vertical lines mark the upper and lower bounds of the highlighted magnitude range that was used to determine  $\bar{\alpha}_{\text{obs}}^{\text{MICE2}}$ . The dashed red horizontal line marks the  $\bar{\alpha}_{\text{obs}}^{\text{MICE2}}$  estimate and the blue dot-dashed horizontal line marks the effective  $\bar{\alpha}_\kappa^{\text{MICE2}}$  determined from the weak lensing convergence with equation (4) and used to calibrate  $\bar{\alpha}_{\text{obs}}^{\text{MICE2}}$ .



**Figure A2.**  $\bar{\alpha}_{\text{obs}}$  estimates from the MICE2 simulations for the magnitude-limited case ( $i < 20.2$ ) over  $i$ -band magnitude ranges below the turn-off magnitude ( $\Delta i$ ) considered to calculate the weighted average. Two redshift bins are considered:  $0.2 < z \leq 0.5$  (top) and  $0.5 < z \leq 0.75$  (bottom). The red cross marks the  $\bar{\alpha}_{\text{obs}}$  estimate that overlaps the most with the  $\bar{\alpha}_\kappa$  estimate from the weak lensing convergence (black line).

less, it still provides a check that allows us to understand how the method estimates  $\alpha$  in the absence of a complex selection function.

In addition, when looking at Fig. A2, one finds that for zlow the  $\bar{\alpha}_{\text{obs}}$  estimates are accurate over a large domain of magnitude ranges (being less than  $1\sigma$  apart when considering  $\Delta i$  anywhere between 0 and  $\sim 2$ ). This confirms that a power law is a good approximation for the luminosity function over a large magnitude range near the faint end of the distribution that implies that, in a magnitude-limited survey,  $\bar{\alpha}_{\text{obs}}$  estimates are robust and accurate even after substantial changes in the magnitude range considered. We find similarly good agreement between the  $\bar{\alpha}_\kappa$  and  $\bar{\alpha}_{\text{obs}}$  in the zhight bin where  $\bar{\alpha}_{\text{obs}}^{\text{zhight}} = 3.12 \pm 0.2$ , while Fig. A2 shows that this estimate is robust at high redshifts.

Despite the consistency between  $\bar{\alpha}_\kappa$  and  $\bar{\alpha}_{\text{obs}}$  and the robustness of the estimate to small changes in the calibration magnitude range  $\Delta i$ , it is surprising to see such a drastic increase in  $\bar{\alpha}_{\text{obs}}$  between zlow and zhight. This seems to be a consequence of the magnitude limit at  $i = 20.2$  being low enough to exclude a substantial fraction of faint galaxies at high redshifts, such that the power law in flux assumed in equation (2) no longer applies. If we consider the luminosity function of the galaxies as a Schechter function (Schechter 1976), such a selection of bright galaxies would lead to a dominant exponential term in the Schechter function that leads to overestimates of  $\alpha$ . In general, this is not of much concern, since most magnitude limited surveys operate within a regime where the power law approximation holds.

This paper has been typeset from a  $\text{\LaTeX}$  file prepared by the author.